

New drug law to speed scientific review

By charging drug companies 'user fees' that will be spent hiring new staff, the US Food and Drug Administration will reduce the time it takes to review new pharmaceuticals from an average of two years to one.

FOR years, the US pharmaceutical industry has complained bitterly that the Food and Drug Administration (FDA) is too slow in reviewing and approving new drugs for the medical market. In turn, FDA has said it does not have enough scientific staff to handle more than 100 applications a year any faster than it does. So some valuable medicines have languished in the bureaucracy for years, unavailable to patients who need them.

The urgent need for rapid review has been nowhere more apparent than in the treatment of AIDS, where true medical need, combined with a strong political lobby of so-called AIDS activists, prevailed to convince the FDA to move quickly on the approval of the drugs AZT and DDI, suspending certain requirements of proved efficacy.

A long overdue bill passed last week by the US Congress should go a long way to right that situation (see page 567). But the legislation breaks new ground by subtly changing the relationship between the pharmaceutical industry and FDA by allowing the agency to charge companies substantial 'user fees' to review a drug application. Fees of \$100,000 per application will apply this year. By 1997, fees will rise to \$233,000 per application. It is the first time the FDA has been authorized to charge fees to commercial interests that, by law, must seek its services.

So the threat of the new law had the pharmaceutical companies howling in protest? Surprisingly, not at all. The industry believes that fees of a hundred thousand dollars or more will be well worthwhile if the time from submission to FDA action can be reduced from two years to one for 'ordinary' drugs, and from one year to six months for potentially life-saving drugs for the terminally ill — AIDS and cancer patients, for instance. In the economics of the drug business, companies stand to make millions of dollars for every month cut from the approval process.

This development represents a change of heart. Although the idea of charging fees has been debated for years, the industry has vigorously protested that fees would be grossly unfair unless FDA spent the money solely on speeding the approval process. The fear was that the money would be dispersed throughout the agency or, worse still, would wind up in the US Treasury to be spent on government programmes in general. Only an agreement reached in the past couple of months to limit user fees to the drug review process has broken the deadlock.

With revenues from fees expected to reach \$330 million within the next five years, FDA Commissioner David Kessler has promised to hire 600 scientists and other professionals to review the safety and efficacy of new drugs. That alone will greatly enhance the overall scientific strength of FDA, which is vital at a time when the number of new drug applications from the biotechnology industry is expected to increase from dozens under review this year to as many as 300 two years from now and, perhaps, as many as 1,200 by the end of the decade.

Two questions remain, of which the more immediate is whether the profits likely to accrue to the pharmaceutical companies will lead to lower drug prices for consumers. Naturally, the companies are promising that faster approval will mean lower prices. They have no choice, having argued in the past that one reason for high prices is that the approval process takes so long. It is tempting, but perhaps naïve, to hope that lower prices at the pharmacy counter will derive from this otherwise sound legislation. Perhaps FDA could use some of its new recruits to keep a close eye on drug prices, using what it learns in renewing the Prescription Drug User Fee Act of 1992, which expires in 1997.

The other question is how far the principle of user fees will spread. FDA is not the only regulatory agency whose applicants complain at the delays in the bureaucracy. People seeking to run television stations or to extract minerals from a mine must also often await permission from the federal government. In its regulatory role, the Environmental Protection Agency is a fruitful source of frustration. It is to be hoped that the new Congress does not take the precedent of the deal with FDA to levy user fees wherever they can be made to stick, which could be unjust. □

Biodiversity supported

Funding for a new programme, with only \$1.5 million for grants, is hardly enough to scratch the surface.

ALMOST all scientists will agree that it is important to maintain biodiversity on this planet, for the good of the planet and the people who inhabit it. It is also perhaps fair to say that biodiversity does not rank high on the list of concerns facing the average citizens of either developed or



developing nations who are all, to one degree or another, worried about more immediate problems — the economy, for instance.

Lack of commitment may account for the fact that the US government has just announced a plan to “encourage the preservation of the world’s disappearing ecosystems and development of new drugs from natural products”, while at the same time committing ridiculously little money to the effort. The new International Cooperative Biodiversity Groups Program, sponsored jointly by the US Agency for International Development, the National Science Foundation and the National Institutes of Health, offers the world a whopping \$1.5 million in grant funds “to support developing nations’ efforts to conserve native species” and enhance research capacity.

Wishful thinking

The biodiversity programme was created following a 1991 conference with the requisite diversity of participants — representatives of government and drug companies, lawyers and ethnobiologists. The presumption is that modest funding from this new US programme will enable researchers in developing countries to collect native species of plants and organisms, train young scientists and create biodiversity research centres. But \$1.5 million will not multiply like the loaves and fishes to feed the need for conservation of species. It is an important goal but one marked by wishful thinking.

The scientists and others at the three agencies who created the new biodiversity program no doubt are keenly disappointed with the funding that has been allocated to such an important project.

In this context it is worth renewed consideration of the role of private industry in species conservation. Merck & Co., one of the world’s largest pharmaceutical companies, has invested \$1 million in a research project in Costa Rica alone to identify new drugs from indigenous plants. Although the terms of that agreement are being questioned (see page 565), the private sector may be better placed to devise ways of protecting species while also using them to human advantage than are governments that fail to back their philosophical commitments with useful sums of money. □

Copernican Columbus?

Columbus, for all the criticism of the past few weeks, deserves his place in history.

POOR Christopher Columbus has had a bad press these past few weeks, being held responsible for all manner of dreadful happenings, from the sharp and cruel attenuation of the native populations of both Americas to the flourishing of the drug trade by symbiosis between rural industrialists in South America and the unhappy populations of cities in the North. So far as is known, Columbus has not been blamed for the parlous condition of public education

in the United States, but nobody would be surprised if he had been. Meanwhile, even the formal celebration in the United States last Monday of the 500th anniversary of his landfall in the Caribbean was overshadowed by the increasingly torrid climate of the presidential election. Why should one who helped to change the world be dealt with so shabbily?

No doubt the respect that would ordinarily be due to a great navigator is compromised by the general knowledge of how Columbus became the one first to reach the New World from the Old in sufficiently good shape as to be able to return. But at a time when seamen around the Mediterranean had developed the techniques of sea navigation to enviable (and often feared) art, Italian Columbus took himself off to rich Spain in search of backers. Despite the passage of five centuries, the story smacks of how itinerant software makers now hawk the contents of their heads between one potential backer and another, settling eventually with the one who offers most. That sounds discreditable.

It also goes without saying that Columbus was lucky to be the first to get to the Americas and to return. Norse and Celtic legends of transatlantic voyages in earlier centuries are entirely believable, but are of no historical significance. They made nobody outstandingly richer than his fellows and, more seriously, they were unrecorded. Columbus, by contrast, entrusted his public relations to the court of Spain. That lesson has not been lost on his successors in the modern age. Do not imagine that when the US National Aeronautics and Space Administration sends its first man or woman to Mars, the rest of us will not be told.

What Columbus’s critics have been overtly complaining about these past few weeks, however, is that he was not required in advance to contemplate the consequences of his voyage; in blithe irresponsibility, he set out without the benefit of even an environmental impact statement, let alone the political, social and economic impact statements expected of one wishing to make such a momentous discovery. But that, of course, simply gives the game away. To Europeans in the fifteenth century, the discovery of the Americas was a true discovery because its scale overwhelmed their capacity to understand it. The cupidity of succeeding decades is a proof of that.

There remains a curious thought about the discovery of the Americas: the voyage was made possible by a primitive skill in the measurement of longitude, the practice of which cannot but have made Columbus and his contemporaries conscious of the Copernican question why the patterns of the supposedly fixed stars should vary with the seasons. Columbus, given his backers, would have known that discretion on the heliocentric issue would be prudent. It is nevertheless remarkable that the whole of modern science postdates his voyage. More than that, merely 15 or 20 generations have passed since then, which is perhaps the most vivid proof there could be of how his discovery accelerated the pace of change. In one sense, it was almost yesterday. □

If biological diversity has a price, who sets it and who should benefit?

Washington. A one-year-old agreement between Merck & Co., Inc. and a group of Costa Rican scientists coupling economic development with the preservation of biodiversity in Costa Rica (see *Nature* 353, 290; 1991) has raised the question of who owns a nation's biological wealth. The agreement, which gives the world's largest pharmaceutical company the right to search for new medicines among the plants, insects and microorganisms collected from Costa Rica's protected forests, is seen either as a novel way to finance biodiversity or as an unconscionable sellout of the country's environmental heritage.

The arrangement has inspired a US legislator to propose an international aid programme to help nonprofit organizations within Latin America and the Caribbean to obtain the necessary scientific knowledge and commercial sophistication to enter into agreements such as that between Merck and the Instituto Nacional de Biodiversidad de Costa Rica (INBio). At the same time, the Costa Rican legislature is expected to vote soon on a bill that would provide increased public scrutiny and a larger share of the proceeds from such arrangements. The legislation has been revised to meet the concerns of drug companies planning to emulate Merck as well as the private research organizations that would receive royalties from the sale of any products developed from indigenous samples.

Much of the opposition to the agreement, led by the Museo Nacional de Costa Rica and a host of nongovernmental environmental organizations within Costa Rica, stems from INBio's private status. Critics are concerned about the confidential nature of such agreements, the lack of public accountability and the failure of such agreements to protect the rights of various social groups within Costa Rica. Mario Carazo, president of the Costa Rican environmental group, AMBIO, argues that although it is called a 'national' biodiversity institute, "INBio has no right and no authority to represent Costa Rica".

Carazo would like to see a larger role for the public in such agreements and a different distribution of the proceeds. Under the terms of the agreement, Merck has agreed to pay INBio \$1 million over two years in return for the opportunity to screen thousands of plant, insect and soil samples for new molecules with biological activity. INBio will receive 5 per cent royalties on the sale of any products developed from these

samples. Costa Rica's Ministry of Natural Resources, Energy and Mines will receive \$100,000 of the \$1 million paid to INBio by Merck and 50 per cent of any royalties — monies that are earmarked for the conservation of biodiversity within Costa Rica. INBio intends to sign similar agreements with two European pharmaceutical companies before the end of the year.

The agreement between Merck and INBio

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

An INBio parataxonomist with a specimen.

treats genetic resources as a natural resource not unlike copper or timber. If that idea is accepted, says Jack Kloppenburg, a rural sociologist at the University of Wisconsin, the question of whether to pay developing nations for access to biological resources becomes instead a matter of "who pays, and how much?"

The environmental organizations do not oppose the commercialization of Costa Rica's natural resources and acknowledge the value of INBio's decade-long effort to create a biological inventory of the country's estimated 500,000 species. They believe, however, that the Costa Rican people should know what biological resources are being sold on their behalf and at what price. Anna Sittenfeld, INBio's director for biodiversity prospecting, says that INBio decided not to divulge such information to prevent organizations in other countries from learning how much Merck was paying per sample and offering Merck the same deal for less than it was paying INBio.

Many believe that the country's best hope for a larger share of the benefits of the country's rich biodiversity lies with a bill, introduced in 1990, declaring all plant and animal species a national patrimony and placing them in the public domain. Organizations would need to obtain special concessions from the government to operate in protected areas.

The bill would weaken INBio's almost exclusive control over commercializing the country's wealth of biodiversity and distributing its benefits. Opponents have removed language that would have prevented companies from patenting inventions arising from the evaluation of Costa Rica's biodiversity. The bill, which covers flora and fauna but does not extend to microorganisms, must be voted upon before the end of December.

The Merck-INBio agreement, praised by such organizations as the World Resources Institute, the US National Academy of Sciences and the Royal Society in Britain, has become a model for the world because it was the first well-publicized deal to involve a developing nation and a multinational company. But critics argue that what may work in Costa Rica, which has a stable democratic government and where one-quarter of the land area has been set aside for the purposes of conservation, may not work so well in other Latin American countries. "INBio can be one model", says Elaine Elisabetsky, a pharmacologist at the Federal University of Rio Grande do Sul in southern Brazil, "but it shouldn't be the model".

Shortly before the US Congress adjourned last week, US Representative Robert Torricelli (Democrat, New Jersey) introduced a bill to set up a Western Hemisphere Biodiversity Cooperation Program within the US Agency for International Development. The programme would help nonprofit organizations within Latin America and the Caribbean to set up biological inventories and databases for purposes of conservation, science or commerce. Entitled the Economic Leadership through Environmental Cooperation Act, the bill is designed to encourage investment, principally by US pharmaceutical companies, in programmes to conserve biodiversity in Latin America.

Torricelli believes that grants of not more than \$100,000 to nonprofit organizations would give them the necessary institutional capacity to negotiate with and enter into bilateral agreements with institutions in the United States. One of the reasons that INBio was such an attractive partner to Merck was because it knew how to process samples and had already conducted an inventory of its enormous biological wealth.

Congress took no action on the bill, but Torricelli says that he hopes to reintroduce it in January with the support of a newly elected Democratic administration more receptive to the issue.

Diane Gershon

New rules loom for US science funding as Congress passes a lean 1993 budget

Washington. The end of the 102nd Congress last week also marks the end of a commitment to a steady growth in the federal funding of science. Next month's election of a US president and a new Congress promises to usher in an era in which research must demonstrate its relevance to society or to an important constituency to win favour with politicians. A close review is under way of a 45-year compact between the government and the academic community that assumes the value of basic research, and the elimination next year of artificial barriers on federal spending promises to intensify competition for scarce funds.

The news for 1993 is not good: the \$1.87-billion core research programme at the National Science Foundation (NSF) will receive \$13 million less than in 1992, for example, and the extramural research budget of the National Institutes of Health (NIH), some three times larger, will barely keep up with inflation. Space science within the National Aeronautics and Space Administration (NASA) is being squeezed to make room within a level budget for the continued construction of space station Freedom and a new advanced solid motor rocket. The Superconducting Super Collider, a presidential favourite, was given only one-fifth of the \$160-million increase it requested, to \$517 million, and the multi-billion-dollar budgets of initiatives that involve several agencies, including global climate change, science education and biotechnology, will be fortunate to stay at current levels.

No reputable analyst will even venture a guess about fiscal year 1994, which begins on 1 October 1993; there are simply too many uncertainties. Besides the results of the forthcoming election, there is much talk of renegotiating the terms of a 1990 budget agreement that would lift next year the limits on domestic discretionary spending (which includes all civilian research) and allow all programmes — military, civilian and international — to compete for the same amount of money, about \$535 billion, that was available in fiscal year 1993. How research would fare in that scramble is anybody's guess. Optimists see science carving out a slice of the military pie, while pessimists see it losing out in competition with jobs, housing, health care and other pressing domestic concerns.

Clearly, it is no longer enough for each agency to propose programmes that its constituents want and then hope that such proposals will be inserted into the budget. The

fight is no longer, if it ever was, between 'big' and 'little' science, or between research and education, or even among disciplines. Bernadine Healy has spent much of her 18 months as NIH director trying to break the biomedical research community of its habit of defining success in terms of new R01 (individual investigator-initiated) grants because that goal is no longer enough to persuade Congress to increase the NIH budget. At NSF, director Walter Massey no longer mentions a presidential promise to

nevertheless, the outlines of this year's federal spending on science are beginning to emerge. Here are some highlights:

National Science Foundation

Although NSF officials expected 1993 to be a difficult year, it was still a shock that Congress refused the entire \$337-million increase they had requested for the six research directorates. (At the same time, Congress added \$8 million to a \$14-million requested increase for science education and

appropriated \$50 million, \$17 million more than requested, for large instrumentation and rebuilding academic facilities.) Congress also said that \$75 million — nearly the full amount requested — must be spent this year on such costly new projects as the Laser Interferometer Gravitational Wave Observatory (LIGO), two 8-metre telescopes and a national high-magnetic field laboratory. The agency dislikes such 'earmarks' because it gives officials less flexibility in deciding which projects will be delayed or trimmed because of a lack of funds.

Language accompanying the spending bill ordered NSF to incorporate the recommendations of its Commission on the Future of NSF (see *Nature* 359, 261; 1992) into its operating plan for 1993

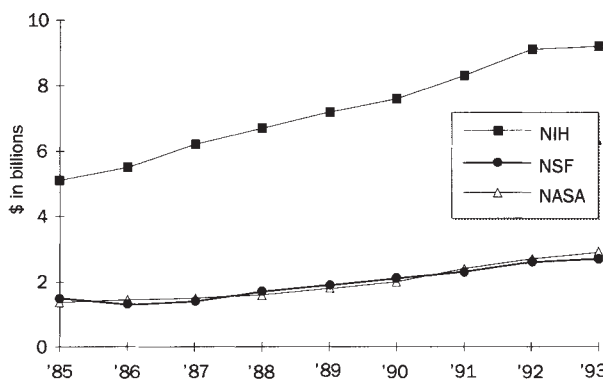
which must be submitted to Congress on 15 December. Although Congress toned down instructions in the Senate version of its appropriation that would have required it to pay more attention to applied research, the final language informs NSF that "it must change the fundamental emphasis it places on various research initiatives"; simply asking for more money for such programmes, it points out, is not the answer.

National Institutes of Health

The NIH budget will rise just three per cent, to \$10.363 billion, a figure that now includes the research components of the three institutes that made up the former Alcohol, Drug Abuse and Mental Health Administration, which merged with NIH in August. With biomedical research inflation somewhere between four and five per cent, NIH's 1993 budget looks like an effective cut, especially compared with last year's increase of nearly \$1 billion.

What this will mean to research grants is an open question. Unlike some previous funding bills, the 1993 budget does not include a specific target of 6,000 new and competing grants. It does, however, include a token \$5 million for laboratory improve-

Growth slows for US research budgets



NIH figures do not include ADAMHA agencies. NASA figures represent space science and applications budget. Source: Office of Management and Budget, Federation of American Scientists.

double the NSF budget within five years, first made in 1987, because he knows that it cuts no political ice.

Why is this? Massey tackled the question in surprisingly frank comments during a meeting last week of the National Science Board, which sets policy for the foundation. "It has become clear that there are major disagreements between our view of things and the view of Congress and the public", he stated. "The public hears that we're No.1 in science, and they want to know why that fact isn't making their lives better. The one thing that works in this country doesn't seem to be paying off, and I don't think that it's enough for us to say that we don't know why and that it's not our problem."

Without a crystal ball to forecast the economy, science policy-makers are left with reality, and they do not much like what they see: for the first time in recent memory, for example, Congress failed to match even the president's modest (an additional 4.5 per cent) request for NIH, and it flatly rejected a proposed 18 per cent increase for NSF's research programmes. The true impact of those meagre budgets will not register with the scientific community until each agency decides precisely how to spend its money;

ment and construction at grant-receiving institutions, a new programme that is being welcomed by the research and university groups that have lobbied for years for NIH to address the problem of crumbling infrastructure at US universities.

In a related matter, Congress decided at the last minute to abandon temporarily a bill that would have offered an alternative to the current administration ban on fetal tissue research and established or modified several other NIH programmes, including those on the health of minorities and women. "We ran out of time, not votes", says Maureen Burns of the Association of American Universities.

National Aeronautics and Space Administration

As with NIH, the 1993 NASA budget will be essentially flat. But space science was treated relatively well. Big research projects such as the Earth Observing System, the Advanced X-Ray Astrophysics Facility and the Cassini mission to Saturn received about as much as the agency had requested, and space science received altogether an increase of slightly more than 4 per cent. Given that the overall NASA budget represents no growth at all — even for inflation — something had to give to make up for space science's rise. The obvious targets were those furthest away: Congress eliminated funding in 1993 for the planned National Aerospace Plane and for missions to the Moon and Mars. Although that action does not kill the two projects, it reduces the chances of success in fiscal year 1994.

For researchers who use satellite data, perhaps the most significant action as Congress moved towards adjournment was a major reform of the Landsat remote-sensing programme (see *Nature* 359, 353; 1992). The legislation will return Landsat to the government, ensuring a much-needed seventh orbiter and lower prices.

Department of Energy

Research spending by the Department of Energy (DOE) will grow hardly at all from 1992 levels. Funding for most programmes rose at most by one or two per cent, and some categories, including both high-energy and nuclear physics, suffered cuts. But Congress did provide the necessary sums for several physics projects under construction, including the Continuous Electron Beam Accelerator Facility (CEBAF) in Newport News, Virginia, the Relativistic Heavy Ion Collider (RHIC) at Brookhaven National Laboratories, and the Advanced Photon Source (APS) at Argonne National Laboratory. However, the Argonne project may yet fall victim to mandatory reductions to be negotiated between DOE and Congress during the next month that would delay by almost a year the opening, now scheduled for spring 1995, of some of its beam lines.

**Christopher Anderson,
Jeffrey Mervis & Traci Watson**

Human Frontiers looks for leader in Europe; US wavers

Tokyo & Munich. In a complex contest of science and politics, three leading scientists from Europe are vying to take charge of the Strasbourg secretariat of the Japanese-inspired Human Frontier Science Program (HFSP), which supports international research on the brain and biological functions. The situation will be further complicated if

ties) may nominate a candidate. Although the official deadline for nominations was 15 September, Japanese government officials have indicated that late entrees will be accepted. The US candidate is Lawrence Bogorad of Harvard University. Bogorad has refused to comment, but knowledgeable sources in Washington confirm that his name is likely to be submitted once bureaucratic hurdles are cleared, perhaps as early as this week.

In an odd twist, Sydney Brenner was nominated not by Britain but by Italy. The United Kingdom did not nominate a candidate because of its displeasure with the treatment given to Gowans, whose name

was put forward at the urging of Japan when the Frontiers organization was established in 1989. Japan did not nominate anyone to avoid appearing to dominate the programme.

The choice of secretary general will be made in Strasbourg next month by the HFSP board of trustees after the nominees have been screened by the HFSP council of scientists. The trustees must decide the extent to which he must be an expert in the fields supported by the programme and whether he should be a day-to-day administrator or more of an intellectual leader. Politics will also play a part: one country has already hinted that it will give more money to HFSP if its candidate is chosen. Japan now provides most of the funds for the nearly \$40-million programme.

David Swinbanks & Alison Abbott



Sydney Brenner, Horst Sund and Michel Cuenod

the United States decides to nominate its own candidate.

The present secretary general, Sir James Gowans of Britain, will step down next spring following differences of opinion with Japanese and some European organizers of the programme over its direction (see *Nature* 358, 527, 1992). So far, the three candidates for secretary general are Sydney Brenner, formerly at the Medical Research Centre and now at the University of Cambridge School of Medicine; Horst Sund of the University of Konstanz, Germany; and Michel Cuenod, director of the Brain Research Institute at the University of Zurich.

Any of the members of the programme (made up of Japan, the United States, Canada, the United Kingdom, France, Germany, Italy, Switzerland and the European Communi-

User fees to hasten FDA review

Washington. Drug companies will begin paying a fee to help the US Food and Drug Administration (FDA) to review new drug applications more quickly under legislation approved last week in the waning days of the 102nd Congress. With the FDA's budget increasing by only 2.6 per cent this year, forcing companies to shoulder some of the financial burden is the only way that FDA can keep pace with its increasing workload, particularly in biotechnology.

The idea of charging users is not new: earlier this year, for example, the administration proposed raising \$200 million of the agency's \$791-million budget with such fees. The drug industry has traditionally opposed the idea, calling it a tax on innovation that would be particularly burdensome

to new, small companies. But it acquiesced after Congress and the FDA agreed that the revenue would supplement — not replace — the agency's regular appropriation.

User fees are expected to bring in about \$330 million during the next five years, allowing the FDA to eliminate its backlog of new drug applications within two years and to shorten the average time of review from 24 to 12 months. To do this, the FDA would hire 600 new reviewers.

The bill takes into account the precarious financial status of small companies by permitting those with fewer than 500 employees and without products on the market to pay only half of the \$100,000 fee for a new drug application and to defer payment for one year.

Diane Gershon

Sandoz-funded research from Rhine spill suggests blockages are worse than toxins

Basel & Munich. A river can quickly recover from 'catastrophic' pollution, even on the scale of the 1986 Rhine disaster, but it cannot recover from engineering projects that obstruct its flow. This unexpected discovery was made possible by a SFr10 million (US\$7.7 million) fund set up by the Swiss chemical company Sandoz as a public apology for causing Europe's worst episode of river pollution.

A fire that destroyed one of its chemical storage depots near the Rhine and caused chemicals such as mercury, disulfon and thiometon to leak into the river damaged the reputation of Sandoz along with the Rhine's ecosystem. Environmental authorities from the four countries through which the Rhine flows — Germany, Switzerland, France and the Netherlands — were hindered by a lack of basic information.

The 'Rhine fund', which has supported 36 independently reviewed projects over a five-year period, was intended to provide a more accurate picture for the future and, although originally castigated as 'guilt money', appears now to have been money well-spent. The work it has funded suggests that the impact of the damage was not as great as predicted.

Researchers found that the colloidal nature of the Rhine's water protected it from the full impact of the poison. The high levels of inorganic particles in the Rhine, constant

throughout the year, adsorbed much of the poison, limiting its effective concentration.

In addition, studies of microinvertebrate populations, indicators of the ability of a river to support life, showed the river's natural ability to recover, provided that it had not been unnaturally engineered. The poison destroyed populations of the microinvertebrates downstream of Basel, but within three years their densities and varieties had returned to pre-accident levels. Conversely, the microinvertebrates failed to return upstream of Basel, an area unaffected by the Sandoz poisoning but obstructed by technical barriers. The obstructed flow caused their habitat, interstitial spaces in the river bed, to become filled with fine mud and decimated the population.

Researchers also found that regeneration of fauna populations comes not from the Rhine itself but from tributaries whose beds act as efficient 'reservoirs of life'. This helps to explain why a section of the Rhine in France whose tributaries were all blocked during the last century has not yet fully recovered.

"While we're not saying that anyone should deliberately pollute rivers, we have always suspected that it would be worse to block a river than to poison it", says Ivan Tomka of the University of Fribourg in Switzerland. But Tomka fears that governments may ignore an important political

question raised by these findings: to what extent should ecological considerations be pitted against the desire of industry and agriculture to tame the river?

Other research supported by the Rhine fund, which was distributed by an independent board of reviewers to scientists in the four countries through which the Rhine flows, revealed a surprising diversity of fish and fauna species. Ecologist Markus Ritter of Basel says that the Sandoz money, small by most standards, brought a "surprising advance in understanding of a much-neglected field."

But the end of support from Sandoz has left some projects in limbo. A plan to establish an international database for Rhine ecology was narrowed to a catalogue of data sources when it became clear that political and legal problems would make it hard to collect the data. The catalogue was compiled two years ago by the Institute for European Environmental Policy, an independent nonprofit organization based in Paris, but has remained uncomputerized and the information is still not freely accessible.

Since the accident, Sandoz has improved its safety standards and increased its investment in environmental protection from SFr66 million in 1987 to more than SFr200 million today. But it has no plans to set up a successor to the Rhine fund.

Oliver Klaffke and Alison Abbott

Science fares well in tight 1993 French budget

Paris. Science emerged relatively unscathed from the austere French budget for 1993 announced earlier this month. The proposed FF53.7 billion (US\$11 billion) budget for civil research and development is up 5.25 per cent more than the overall growth of 3.4 per cent in public expenditure. For the second year in a row, research organizations also are virtually the only public bodies spared from retroactive reductions.

Why is that? Research has been a national priority for the Socialist government since it came to power in 1981, and funding for research has grown from 1.97 per cent of gross national product (GNP) in 1981 to 2.4 per cent in 1990. Although the government has fallen short of its target of 3 per cent of GNP, the growth has been enough to keep France in fourth place worldwide behind Japan, the United States and Germany.

Much of the credit goes to Hubert Curien, whose ministry of research and space has developed a coherent long-term national policy, recruited talent, prodded research

organizations into becoming more efficient and encouraged industry to carry out more research through subsidies, tax credits and the like. In turn, industry has spent a record amount on research for the past four years, some 1.1 per cent of GNP, at the same time reducing advertising budgets by 30 per cent.

This year's budget continues the themes of past years, although support for industrial research has been somewhat curtailed. The Centre Nationale de la Recherche Scientifique (CNRS), the country's largest research organization, must make do with a proposed increase of only 4.5 per cent, to FF12.4 billion, compared with 8.1 per cent last year. In contrast, the budgets of the space research organization (CNES) and the medical research organization (INSERM) are scheduled to be increased by 7.5 per cent and 6.6 per cent.

The French atomic research agency, which has had its budget reduced steadily over the past few years, received a proposed

increase of 3.8 per cent to FF 6.4 billion. Ministry officials feel that the agency has undergone sufficient reforms to merit renewed support. The decision to reinvest is linked to an increased emphasis on safety, cleaner fuels and the technological challenges of dismantling reactors while research into plant physiology will be phased out during the next three years and work on protein engineering will continue only as collaborative projects with other organizations.

Among the smaller organizations, the National AIDS agency received a proposed increase of 10.5 per cent in its budget, while the Centre d'Etude du Polymorphisme Humain (CEPH) will have its budget increased by 40 per cent.

The creation of scientific jobs is another priority for the French government. Next year's budget promises to add 470 jobs, bringing to 1,625 the number of new positions created since 1988.

Declan Butler

Britain wrestles with EC rule on modified organisms

London. British lawmakers are trying to incorporate both a desire for openness and the preservation of patent rights into an interpretation of two directives from the European Commission on genetically modified organisms. Biotechnology companies are concerned that the directives, which were supposed to become law a year ago, will discourage research and compel the release of information that could invalidate patent applications.

The directives, which cover the contained use and release of these organisms, say that the commercial value and intellectual property of researchers will be preserved but that information about the organism and its proposed use must be circulated within the European Communities and made public in the host country, probably through accessible registers. The only permanent solution would be to change the directives, a time-consuming process involving considerable loss of face on the part of the commission and one that seems unlikely in the current political climate. Another option is to change the patent laws, but this would require agreement by each member of the EC. Several countries, including France and the Netherlands, intend to implement regulations that follow the

directive, while Denmark makes the information available already.

Although the regulations were revised earlier this year after lobbying by the biotechnology community, the current draft has retained strict requirements about disclosure. Industry is concerned that the rules would lead companies to reduce or abandon promising research or force them to shift operations to countries without such rules. University researchers also fear losing intellectual property rights over their work.

In the coming weeks, the Department of the Environment (DOE) and the Health and Safety Executive (HSE) will try to strike a balance between protecting property rights and ensuring the public's right to know about such experiments. It is hoped that the regulations will be ready for Parliament by the end of the year.

One solution is for researchers to submit two sets of information, one giving full details of the work on which the DOE and HSE could grant or deny permission to go ahead, and another for inclusion on public registers. However, the balance between the two sets would be delicate, and researchers may not want to wait until their patent is rejected to learn whether they have given too much away.

Ian Mundell

Overregulation could damage US biotechnology, says report

San Francisco. Existing US regulations for field tests of genetically engineered organisms are stifling research, according to a new report by the National Biotechnology Policy Board that recommends strict regulation of deliberate releases only for products posing the greatest risk to public safety.

The document emphasizes the social costs of regulatory delays and recommends a pilot study to measure the benefits of new products. This practice would help to quantify the impact of regulatory delays, provide data about the safety of new products, educate the public about the potential of biotechnology and reward federal agencies for prompt review, says the report, which was approved last week by the policy board at its meeting in Washington, DC. The board is a 20-member body formed by the National Institutes of Health in 1989 at the direction of the US Congress to review government activities relating to biotechnology.

The board's recommendations follow a study by the congressional Office of Technology Assessment that warned regulators not to relax biotechnology regulation so

much that public concerns are ignored. But the report points out that "intense government oversight tends to confirm public perceptions that biotechnology processes pose significant and unique dangers".

The head of the biotechnology programme at the US Department of Agriculture (USDA) criticized the report for emphasizing the negative aspects of regulations. "I believe that good regulations not only assure safety but also help to facilitate technology transfer and development", says the USDA's Terry Medley, whose office is expected shortly to release its proposal for somewhat relaxed field-testing rules. At the same time, microbiologist Rita Colwell, president of the Maryland Biotechnology Institute in Baltimore, defended the report for endorsing the importance of field testing.

The report also recommends financial incentives for the industry, including government participation in joint ventures, technology transfer from universities with patent rights attached in some cases and changes in the US tax code that would encourage investment in research.

Sally Lehrman

Brian Henderson named president of Salk Institute

Washington. Brian Henderson, a 55-year-old cancer epidemiologist and director of the Kenneth Norris Jr Comprehensive Cancer Center at the University of Southern California (USC), has been appointed president of the Salk Institute, with effect from February 1993. The announcement, to be made this week, marks the end of a debilitating four-year international search for a successor to Frederic de Hoffmann as head of the private institute, founded in 1960 by Jonas Salk as a centre for basic research relevant to human disease.



Brian Henderson

In choosing Henderson, the institute sought a president who could work well with the institute's 225 researchers, increase its \$27-million endowment (relatively small for an institution with an annual budget of \$40 million) and oversee growth in faculty after a sizeable expansion of the institute is completed over the next few years. It also wanted someone to educate the community about the importance of basic science "so people will see that the work we're doing here affects health", says Stephen Heinemann, faculty chairman.

Renato Dulbecco, who was appointed acting president and then actual president after de Hoffmann resigned in 1988 because of illness (he died the following year), has made some overtures to the community, but de Hoffmann did little to publicize the institute's mission during his 18-year tenure. Henderson hopes that references to his own work on genetic susceptibility to cancer will help to demonstrate the connection between basic research and health care, and he intends to continue an active research programme in epidemiology.

In the past few years, Salk has suffered the embarrassment of failed negotiations on the presidency with such well-known researchers as James Darnell of Rockefeller University and Arnold Levine of Columbia University, leading to public speculation that the institute's trustees were not equal to the task. In response, John Henry Felix, chairman of Salk's board of trustees, assumed responsibility for contract negotiations, a task that had previously been performed by a committee.

Traci Watson

Nobel Prize given for work on protein phosphorylation

Two US biochemists have won the 1992 Nobel Prize in Physiology or Medicine for discovering a fundamental mechanism that regulates cell metabolism. In a series of experiments starting in the 1950s, Edwin Krebs and Edmond Fischer, both of the University of Washington, Seattle, characterized protein kinases, a class of enzyme



Edwin Krebs

that change proteins from their inactive to their active form by speeding the chemical bonding of a phosphate group to the protein. This transfer, or phosphorylation, is the underlying 'life switch' that starts and stops a variety of cell functions, from the breakdown of fats

to the generation of chemical energy, in response to hormonal and other signals.

Working on muscle tissue, Krebs (who is unrelated to Hans Krebs, another Nobel laureate who described in the 1930s the citric acid cycle of metabolism that bears his name) and Fischer showed that protein kinases regulate phosphorylation by catalyzing a reaction in which a phosphate group is transferred from the energy-carrying compound ATP to a protein, thereby activating the protein and contracting the muscle. Hundreds of new protein kinases have since been found, and researchers esti-

mate that perhaps one per cent of the genes in the human genome code for protein kinases of one sort or another.

More recently, the two scientists and their colleagues have identified some of the key enzymes — called phosphatases — that catalyze the transfer of energy in the other direction by removing phosphate groups from proteins. Combined, this mechanism for transferring phosphate groups to and from a protein to turn it on or off is known as reversible protein phosphorylation, and it is for the characterization of this fundamental process involving both kinases and phosphatases that the Nobel Prize committee is honouring the two biochemists.

Krebs won a Lasker award in 1989 for his research; both men are members of the US National Academy of Sciences. About one-third (46 in all) of Lasker winners go on to win a Nobel Prize, making it one of the best predictors of future Nobel laureates. At the time Krebs and Fischer did their work, researchers did not appreciate the significance of protein kinases as a regulatory mechanism.

Krebs, 74, and Fischer, 72, are still active researchers. In 1988, Fischer and his group proved that surface proteins of CD45 T-cells work on the same principle as a phosphatase, removing phosphate groups from an enzyme and starting a cascade in which other enzymes (including some detected by Krebs) are activated. This insight into the human immune system has helped researchers to learn how immune-suppressing drugs prevent the body from rejecting organ transplants.

Christopher Anderson

NOAA appoints ombudsman to repair frayed ties with universities

Washington. Hoping to mend its tattered reputation among academic researchers, the US National Oceanic and Atmospheric Administration (NOAA) has turned to an old standby: the complaints desk. Last month, William Hooke, the deputy chief scientist at the agency, was appointed to the new position of ombudsman and given a mandate to untangle red tape and otherwise help scientists to deal with unresponsive officials and waylaid grants from the agency's \$100-million extramural research programme.

In his first few weeks, Hooke, a veteran NOAA official with a good reputation among academic scientists, has already managed to release a handful of grants that were held up by the bureaucracy. And so far, Hooke says,

he has been able to resolve most complaints within 24 hours.

Academic scientists who have seen other NOAA reforms founder say that it is too early to pass judgement on the latest effort. But Christopher Mooers, a University of Miami oceanographer, believes that an ombudsman may help NOAA to solve its frustrating tendency to allow worthy grant proposals to become bogged down in a patchwork administrative structure that requires several offices to give their approval. An ombudsman may not make NOAA run on a clock more conducive to an academic researcher, but for the first time the agency has designated someone whose job it is to argue on their behalf.

Christopher Anderson

Perkin-Elmer to buy Applied Biosystems to broaden markets

Washington. Applied Biosystems, Inc. of Foster City, California, a leading supplier of automated DNA and protein systems for life-science research, announced last week that it is to become part of Perkin-Elmer Corporation, a manufacturer of analytical instrumentation based in Norwalk, Connecticut.

The deal, which is valued at about \$330 million, brings together two companies with closely related technologies. Perkin-Elmer acquired exclusive worldwide rights to market Hoffmann-La Roche's polymerase chain reaction (PCR) technology for nondiagnostic purposes after Hoffmann-La Roche bought the technology from Cetus Corporation in July 1991 for \$300 million (see *Nature* 352, 364; 1991).

The PCR technology goes "hand-in-hand with our DNA analysis and synthesis technology", says André Marion, chairman and chief executive officer of Applied Biosystems.

Applied Biosystems was founded in 1981 by Marion and Sam Eletre and employs 1,200 people. It reported total annual sales this year of \$183 million compared with \$164 million last year, although the company reported a net loss in both years.

Sales of products for DNA research increased 31 per cent over 1991 levels and now account for more than 60 per cent of total sales. However, sales of the company's protein products declined, reflecting an industry-wide slump. Perkin-Elmer was founded in 1937, employs 6,000 people worldwide and earned \$58 million this year on annual sales of \$911 million.

The move will enable Applied Biosystems to tap into emerging markets such as agriculture, animal husbandry, environmental testing, forensic science and paternity testing. In total, these new markets are larger than the research market that the company now serves.

Under the terms of the stock swap, Perkin-Elmer will exchange 0.678 shares of its stock for each share of Applied Biosystems' stock. The day after the announcement, Perkin-Elmer's share price closed down \$1 1/2 at \$29 1/2; Applied Biosystems was up \$2 3/8 at \$19 3/8.

The newly formed group will be called the Applied Biosystems Life Science Division of Perkin-Elmer and will assume overall responsibility for Perkin-Elmer's life science business, including its PCR amplification technology. The acquisition must still be approved by shareholders and various regulatory authorities but is expected to be completed early next year.

Diane Gershon

Plight of Bosnia and Croatia

SIR — These are great times for the revival and advancement of the theory of symmetry of culpability. I refer, of course, to the war in Croatia and Bosnia, and your leading article (*Nature* 358, 439; 1992), and the News story by Alison Abbott (358, 360; 1992). Here are some facts which would greatly improve the foundations of this theory and, accordingly, the chances for peace in Europe.

Serbs, 12 per cent of the population of the Republic of Croatia, have occupied 25 per cent of its total territory, and effectively 'cleansed' it of Croats, Hungarians, Czechs, Slovaks and Ukrainians. Serbs, 34 per cent of the population of Bosnia and Herzegovina have occupied 60 per cent of the total territory of that state, and are effectively cleansing it of Croats and Muslims. The problem is not agricultural 'land grabbing' (agriculture has never been a favourite subject for communists), but the Serbian quest for control of communication lines and corridors. The loss of these would render both states, Croatia and Bosnia-Herzegovina, a joke in terms of contemporary non-agricultural economy.

These remarkable Serbian achievements are a product of other initial and consequential symmetries: Fighter-bombers: Serbs 600, Croats none; Tanks: Serbs 1,800, Croats none; Heavy artillery: Serbs 2,000, Croats none; Ordnance: Serbs huge stockpiles of the former Yugoslav Army and supplies through Rumania; Croats negligible initially, now carefully dosed life-line supplies breaking the UN embargo through Slovenia and Hungary; Major damaged cities: Serbia none; Croatia 10 (with a total population in excess of 400,000; Bosnia 8 (population in excess of 600,000); Ravaged cities: Serbia none; Croatia 2 (Vukovar and Petrinja).

Casualties are mounting, but for each dead Serb there are 5 dead Croats and 20 dead Muslims (most of them civilians, women and children).

The symmetry in the domain of science should also be mentioned. Croatian scientists share the Serbian fear that "prolonged sanctions will destroy" science. Croatia, with 20 per cent of the total population of the former Yugoslavia contained some 18 per cent of registered scientists (Croatia did not need this war to apply evaluation by peer review, as Glišin hopes for Serbia, 358, 361; 1992). Croatia was forced to contribute 28 per cent of the total Yugoslav federal budget and 40 per cent of foreign hard currency earnings (1986). Croatian scientists received about 10 per cent of federal funds (1987) for research and development, yet produced up to 40 per cent of papers

from the former Yugoslavia cited in the Institute for Scientific Information *Science Citation Index (SCI)*, and issued the only two Yugoslav scientific journals recognized by *SCI* (1985): *Croatica Chemica Acta* and *Periodicum Biologorum*. The cut-off of US sponsored cooperative projects and those with the European Communities (EC) is symmetrical again. Croatia has not been admitted to the EC's PHARE programme.

The theory of symmetry of culpability should also take into account the case of the Interuniversity Centre for Postgraduate Studies in Dubrovnik, a cooperative venture of 250 universities worldwide and year-round courses. As its director, Professor Kathleen Wilkes of the University of Oxford, witnessed, its building took a few well-aimed Serbian shells and burned down along with its specialized library of 25,000 volumes.

For those unfamiliar with the theory of symmetry of culpability, it was originated by Neville Chamberlain and published in Munich in 1938.

Velimir Pravdić

*Ruder Bošković Institute,
PO Box 1016, Bijenička 54,
41001 Zagreb, Croatia*

SIR — To write about difficulties of Serbian science and researchers and at the same time not even mention the situation of science in Bosnia, where dozens of university buildings, research institutions and libraries have been set ablaze or demolished by Serbian mortar fire, is, to say the least, hypocritical. How those burnt and demolished building compare with a \$65,000 computer? How do burnt libraries compare with the Institute of Physics receiving only 40 of its 180 subscriptions? Do you remember the scenes shown on television when a crowd of Sarajevans lining up for bread was hit by a Serbian shell? Do you remember the man lying in the puddle of blood, crying for help and stretching his hands toward the camera? That was Professor Mahmud Dikić, with whom I worked at the Mechanical Engineering Department of the University of Sarajevo, Bosnia and Herzegovina. In that attack he lost his legs. How does that compare with BITNET electronic mail lines being disconnected?

I assume that your aim was to inform us about the protests that the Serbian scientific community is trying to articulate against their government. The protest could be summarized by saying that, after the international community imposed sanctions, 70,000 students occupied the university's main building for 26 days and that all examinations were suspended until late August. What

Abbott failed to say is that this is too little, too late. She also failed to say that the Serbian scientific community is not as innocent as she suggests. There was no mention at all of the role of the Serbian Academy of Sciences and Arts and of its Memorandum (1986) in laying the ideological framework for the formation of Greater Serbia and the atrocities that have followed.

The international scientific community used to be very vocal against the suppression of human rights in Eastern Europe, but now seems to be untouched by the plight of the Bosnian people, human rights abuses and genocide. How many other nations will be put into the Serbian cleansing machine before the world's scientific community reacts?

Sead Dorić

*Institut National d'Optique,
369 Rue Franquet,
Sainte-Foy, Quebec, Canada G1P 4N8*

SIR — When black pictures of history are once again emerging in Bosnia and Herzegovina, where more than 100 detention or death camps for Croats and Muslims have been established by Serbians, I would have expected Serbian scientists to be concerned about their colleagues and collaborators in Bosnia and Herzegovina. But the major worry of Serbians seems to be about shortage of funding, chemicals and about a computer that has been paid for and cannot be imported to Serbia.

Which is the guilty party? The United Nations (UN) for striking back with sanctions that will affect not only Serbian researchers but also the authoritarian communist government of Serbia, or the latter for eliciting such a response from the rest of the world?

Every scientist in Serbia should understand that silence is sometimes the same as a lie or crime especially when it is expressed in such a selfish manner. "As scientists we can only protest we cannot take up guns and kill people" said Dragan Vuckovic, a Serbian scientist. What does he think about Bosnian scientists, Croats, Muslims and loyal Serbs? Are they running their experiments? No, they are not, they are fighting for their own lives and the lives of their children or starving to death in the detention camps. No change is conceivable until human rights are acknowledged in Bosnia and Herzegovina, and until they are willing to help their colleagues to stop the bloodshed.

The UN must exert a stronger embargo on Serbia, in the hope that in the near future science would be restarted in a fair and democratic environment.

Ivan Dikić

*NYU Medical Center,
550 First Avenue,
New York, New York 10016, USA*

HIV: to vaccinate or not to vaccinate?

Gordon Ada, Bob Blanden and Arno Mullbacher

"Improbability of effective vaccination against human immunodeficiency virus . . .", declares the title of a new paper by Dr Albert Sabin. But three immunologists see flaws in his argument.

THE task of developing a vaccine against the human immunodeficiency virus (HIV) is both urgent and difficult. Is it likely to succeed? Dr Albert Sabin, an eminent virologist, thinks not, and last month¹ set out his reasons for that belief. He was quite right to do so, for he raises some serious scientific issues. But in our view such pessimism is premature, and his doubts can be largely allayed on immunological grounds.

Last year, Sabin pointed out² that one potential difficulty was whether any vaccine against HIV would protect a person from infected cells in semen. In his latest paper¹ (the full title of which is "Improbability of effective vaccination against human immunodeficiency virus because of its intracellular transmission and rectal port of entry") he reiterates this message; he quotes several unsuccessful protection experiments with cells infected with simian immunodeficiency virus (SIV) and comments upon an experiment in which vaccinated chimpanzees successfully withstood a challenge of HIV-infected blood cells³. He now draws attention to the high content of virus and of HIV-infected cells in some samples of semen. But in particular Sabin has reservations whether vaccination could be effective against cells containing "cytoplasmic pro-virus" and/or chromosomally integrated HIV complementary DNA, because it was thought that such cells would not express viral products at their surface.

In the vast majority of cases, however, the major histocompatibility complex (MHC) antigens at the cell surface would be foreign (allogeneic) to the recipient. Some of the lympho-myeloid cells would be quite effective as stimulators of a rapid host-versus-graft response, whether or not they were infected, and the rejection (destruction) of all semen-derived cells would follow. One way in which a response by the normal host to allogeneic lympho-myeloid cells could be avoided would be a drastic down-regulation of MHC antigen expression or other co-stimulatory activities in infected cells. This does not occur to a substantial extent in productively HIV-infected lymphocytes or monocytes, and so is even less likely to occur in nonproductively infected cells. Semen is also reported to have immunosuppressive properties, but these would need to be very powerful to prevent an

allogeneic reaction *in vivo*.

Cells containing endogenous C-type viruses have been shown, perhaps unexpectedly, to express viral peptides associated with MHC molecules⁴. By analogy, if the HIV cDNA in live or dead semen cells were to be transferred undamaged in some novel way to live host cells, as Sabin suggests, expression of viral peptides could also occur. If so, the 'transfected' host cells would be recognized by cytotoxic T lymphocytes (CTLs) in the vaccinated host. Sabin rightly emphasizes the importance of cell-mediated responses in controlling viral replication.

There are other possibilities. Might some cells in the semen escape immunosurveillance by quickly reaching immunoprivileged sites in the recipient? We are unaware of any evidence that this is likely to occur. Might a small proportion of the donor's infected cells initiate a graft-versus-host response, so becoming activated and expressing viral antigens? Anti-HIV antibody generated following vaccination should deal with such cells, either by antibody-dependent cellular cytotoxicity or C-dependent lysis. Cells in the donated semen already expressing viral antigen could fuse with host cells (gp120/CD4 interaction), and thus be recognized by both class I MHC-restricted CTLs resulting from the vaccination and alloreactive CTLs. Although there is uncertainty about the likelihood of some theoretical possibilities, we believe that these particular concerns are not a serious barrier to vaccine development and use.

Sabin puts stress on another point, the hazards of receptive anal intercourse, though the extent to which these hazards also apply in heterosexual vaginal intercourse is not clear. The content of infectious virus-infected cells in the semen of asymptomatic, infected men varies greatly, from high levels (quoted by Sabin) to levels undetected by the polymerase chain reaction⁵. In the absence of other sexually transmitted diseases, the efficacy of transmission of HIV infection by sexual intercourse, including receptive anal intercourse, and needle sharing, is very low — 1 per cent or less⁶. Such values, together with a related finding (that the initial viraemia occurring after infection seems usually to reflect the replication of virus of a single antigenic specificity^{7,8}) are consistent with infection being initiated by a low dose of

infectious virus in many cases.

A further issue raised by Sabin¹ is the failure of antibody to neutralize virus during an infection. However, the very fact that escape mutants are found so frequently during an infection indicates that antibody of the appropriate specificity is effective. This view is supported by the finding⁹ that a monoclonal antibody against the HIV gp120 V3 loop protected two chimpanzees from a large dose (75 chimpanzee infectious doses) of virus whether the antibody was administered before or after the challenge. Unquestionably, more needs to be found out about the effectiveness of secretory immunoglobulin A in preventing infection and the feasibility of inducing a persistent response in the rectum and in the female reproductive tract. There is a recent report that monkeys previously vaccinated by parenteral administration against SIV, withstood first a challenge by SIV given parenterally and then a subsequent challenge of SIV delivered via the rectum¹⁰.

There is a long way to go before highly effective anti-HIV vaccines are formulated. Yet, there is reason to believe that 'first-generation' vaccines, even though they may not possess all of the desirable properties, could prevent infection in a proportion of recipients, and much could be learnt to facilitate future vaccine development. An answer will not be available until the best candidates or combinations of candidates, either already under phase 1 trials or about to undergo such trials, meet certain guidelines (including safety) and can then be tested in appropriately designed and executed phase 3 trials. To delay or not perform such trials for the reasons proposed by Dr Sabin would be disastrous for the increasing numbers of people exposed to the risk of infection. □

Gordon Ada, Bob Blanden and Arno Mullbacher are in The John Curtin School of Medical Research, Australian National University, Canberra, ACT 2601, Australia.

1. Sabin, A. B. *Proc. natn. Acad. Sci. U.S.A.* **89**, 8852–8855 (1992).
2. Sabin, A. B. *Science* **251**, 1161 (1991).
3. Fulz, P. N. et al. *Science* **256**, 1687–1690 (1992).
4. Janeway, C. A. 8th Int. Cong. Immun., Budapest, 1992.
5. Krieger, J. N. et al. *J. inf. Dis.* **163**, 386–388 (1992).
6. Blattner, W. A. *FASEB J.* **5**, 2340–2348 (1991).
7. Wolfs, T. F. W. et al. *Virology* **185**, 195–205 (1991).
8. Wolfs, T. F. W. et al. *Virology* **189**, 103–110 (1992).
9. Emini, E. A. et al. *Nature* **355**, 728–730 (1992).
10. Cranage, M. P. et al. *Lancet* **339**, 273–274 (1992).

SCIENCE IN JAPAN

The accompanying article includes the following recommendations for the reform of Japan's science:

- The status of Japan's commitment to double government spending on research and development "as soon as possible" should be clarified, not only in terms of timescale but by agreement on where growth is most needed.
- The administration of academic research grants should be transferred to a new independent research foundation, which should also be responsible for supporting (and sometimes closing) university departments and research institutes. Where necessary, the foundation should use overseas as well as Japanese referees. A threefold increase of present spending is probably required. Grants should be larger and should include larger overheads (say 20 per cent) so departments have an incentive to attract good researchers.
- A group of Japan's national universities, while still supported by public funds, should become self-governing, responsible for their own administrative as well as academic affairs. Otherwise, Japan will not be able to create an internationally outstanding institution of scholarship and research.
- The government needs better machinery for setting and monitoring government science policy. The Council for Science and Technology is an awkward blend of cabinet committee and advisory council whose advice is predetermined by what is politically feasible; at the least, it could be divided into two.
- Geography and language are still serious barriers to Japan's full participation in international science. More travel funds administered by department and laboratory heads, especially for younger people and for short overseas visits, are much needed. Overseas scientists should be employed at public institutions on terms identical with those offered to Japanese.
- Japan should not contribute to the US Superconducting Super-Collider project or, if the political pressure proves irresistible, should find the funds from its substantial ODA (Overseas Development Assistance) budget.
- The possibility of establishing a large industrially supported research foundation should be explored.
- The habitual politeness of Japanese scientists and the tendency not to criticise in public others in the "same village" of research must be attenuated, both internally (in matters such as peer-review) and in the deference shown to pundits from overseas (*Nature*, if need be, included).

Reforming Japan's science for the next century

Japan's science is at a turning point. Japan faces its worst recession in decades and yet the government plans to double spending on research. It is a precarious time, but a golden opportunity for change.

Japan is always changing and yet is always the same. The streets of Tokyo are more crammed than ever with traffic, but the passengers may be women wearing kimonos, on the way to a wedding perhaps. Government offices are mostly air-conditioned now, but serried ranks of bureaucrats still sit as in schoolrooms, within earshot of each other and facing the head of the office and his deputies, with their backs against one wall. In the university laboratories, the equipment is now usually as up to date as anybody could ask, but the old grumbles that *Nature* first noted in its survey of science in Japan in 1972 persist, the grievous shortage of technical assistance, for example.

The coexistence of old and new is a Japanese phenomenon widely remarked upon. The Japanese seem to relish the contrast, typified by the television receivers on the bare boards of the rented

rooms of the traditional Japanese inn, where nothing else rises above knee-level. But the contrast can also be irksome, and nowhere more so than in the administration of science, where the tradition and the bureaucracy remain stifling influences, to the frustration of researchers and to the detriment of their research.

What follows is not so much a formal survey of the present state of science in Japan, extending those of 1972 and 1983, as a manifesto for reform.

We offer a recipe for further change that would remove some persisting impediments to research in Japan, enabling Japanese research to claim the place in the international research enterprise to which its quality and distinctiveness surely entitle it. The manifesto derives from frank conversations with more than 60 Japanese academics, industrial scientists and administrators, some of whose ▷

Nature's roots in Japan

Nature has been reporting on Japanese science since soon after the journal's birth in 1869, also the second year of the Meiji restoration that brought Western science and technology to Japan.

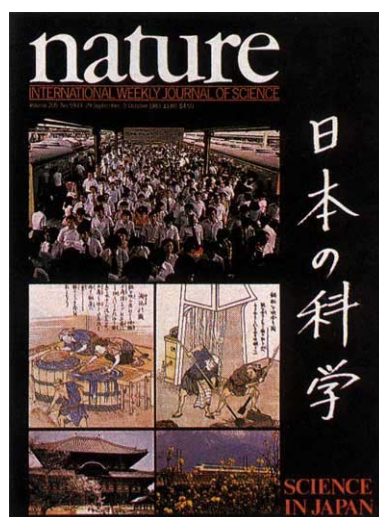
The first reports came during the 1870s from European scientists and engineers who came to help set up universities and research institutions in Japan. The first scientific articles from Japanese authors followed soon

after in the 1880s. For example, in 1887 S. Sekiya, a seismologist at Tokyo University, described in detail the effects of an earthquake that occurred in January that year and warned prophetically of the dangers of "the rapidly increasing use of kerosene lamps" (in 1923, much of Tokyo was destroyed when kerosene stoves toppled by a major

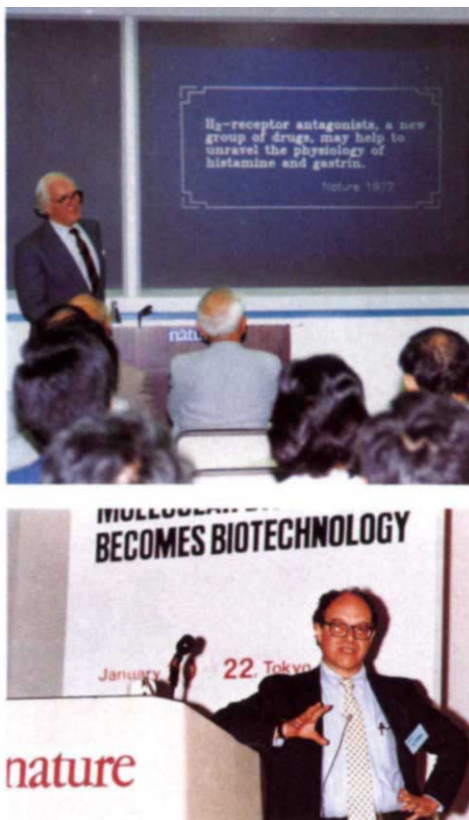
earthquake sent a firestorm raging through the capital).

At the turn of the century, Kumagusu Minakata, a biologist and colourful eccentric intellectual of the Meiji Era who was a friend of the late Shōwa Emperor, became a prolific writer for *Nature*, contributing some 50 letters on subjects ranging from "the constellations of the Far East" (1893) to "colours of plasmodia of some Mycetozoa" (1912). Then, in the 1920s, another

regular Japanese contributor was Hantaro Nagaoka, who was first to propose the "Saturnian" model of the atom, in which electrons revolve around a central proton (although he won little recognition in the West for doing so). Later, Hideki Yukawa published some of his research on mesons in *Nature* which won him the Nobel prize. ▶



Nature's special issue of 1983.



Growing audience

Nature's conferences in Japan, studied with presentations by Nobel prizewinners and Nobel prizewinners-to-be, have proved very popular with Japan's young scientists. The January 1986 conference, entitled "Molecular Biology Becomes Biotechnology", drew an attendance of nearly 500 scientists, half of them under 35 years of age, with presentations from about twenty speakers, including 1981 Nobel prize winner Wally Gilbert, and Susumi Tonegawa, who won his Nobel prize in physiology and medicine a year later. *Nature's* 1988 conference "Horizons in Molecular Biology" drew an even larger audience of young researchers under 35, who came to see speakers such as Harmut Michel, winner a few months later of the Nobel prize in chemistry, and the conference buffet was graced by the presence of the Emperor and Empress of Japan (then Crown Prince and Princess). Earlier this year, *Nature* attracted a similarly large turnout for its first conference in the physical sciences entitled "Nanotechnology: Science at the Atomic Scale", in which Nobel prizewinner Heinrich Rohrer played a key role. And in 1989, *Nature* staged a lecture for pharmacologists and biologists in Tokyo by Sir James Black, who won the 1988 Nobel prize in physiology and medicine.

► ROOTS

More recently, *Nature's* editorial staff have been actively reporting on Japan. In 1972, one of the present authors (J.M.) visited Japan and compiled a 34-page special issue with contributions from many leading Japanese scientists.

It is remarkable that many of the criticisms of Japan's government research, particularly in the universities, raised 20 years ago in that special issue still hold true today, for example, the overbearing administration by Monbusho, the lack of technicians, inbreeding of university researchers and the defects of Monbusho's grant system.

The concept of university autonomy was also discussed briefly in the 1972 issue, but it seems that this Western ideal was no better understood then in Japan than today. Haruo Katsunuma, for example, then dean of the medical faculty of Tokyo University, claimed that "university autonomy ... has been acknowledged ... as a fundamental tenet". The idea that full autonomy requires full administrative autonomy, as well as academic autonomy, was apparently as unrecognized then as now. *Nature's* editor also noted the lack of any equivalent to Western research grant-awarding organizations, such as the United Kingdom's Science and Engineering Research Council.

During the 1980s, *Nature's* place in Japanese science was strengthened enormously. Following a second special issue in 1983, an editorial office was established in Tokyo in 1984. This was followed in 1987 by the set-up of Nature Japan KK, the company that handles all aspects of *Nature's* operations in Japan, and the start of printing of the journal in

NATURE VOL 240 NOVEMBER 24 1992

SURVEY OF JAPANESE SCIENCE

Prospects for Science in Japan

One of the reasons why the organizations and conduct of science in Japan attract the interest of people elsewhere is that, at least from the outside, it seems as if the Japanese government has managed to solve the problem which preoccupies governments elsewhere in the world: how best to ensure that funds invested in science and technology will be economically and socially beneficial. For it is not now plain that in the past few years the economic success of Japan as an exporter has depended on an increasingly efficient application of comparatively pedestrian technology to the manufacture of products now widely recognized to be technologically superior and cheap in well-trodden ways, for example. And outside Japan, it is no apparent that these marvels are not accompanied by the pouring of low wages, so far as comparisons can be made between communities to different, personal incomes in Japan are not much different from those in most of Western Europe even if academics have cause for complaint—that outsiders have been driven to the belief that there is something special, valuable and desirable about the utilization of science for economic ends in Japan.

In the articles which follow, there is much to make more rounded this simple external impression. Mr Y. Nakase, the most powerful influence on the administration of science and technology in Japan, is part because he is Minister of State for Science and Technology in Mr. Tanaka's new government, but also because he is, as well, the head of the Ministry of International Trade and Industry, has very little to say about television sets, indeed, he is much more concerned with the way in which the scientific enterprise can be used domestically in Japan. He writes of what the Americans call health care delivery, the need to improve and even understand the environment and the need somehow to be able to predict the consequences of new technology. To the extent that Mr Nakase's ideas are different from and even orthogonal to those of many of his predecessors, this is a reflection of how the government has come into power less than six months ago in unopposedly uncontested manner. The air route between Tokyo and Tokyo must now be absent as busy as that across the North Atlantic, which is a striking measure of how circumstances have changed. Some might say that Japan is now strong enough to have a foreign policy of its own but what Mr Nakase says is that the time has also come when it should be possible to harness science and technology to the needs of people who live there, not simply to the demands of the markets overseas.

There is no reason to think that these new ideas will be frustrated—on the contrary, if the government of Japan is able to turn its technological energies towards the needs of its own people, it will find that it has done much to ease some of the political conflicts within Japanese society, especially the alienation of the extreme left from the government process.

What people elsewhere could ask themselves is whether the speed with which it has been possible to recent progress of manufacturing industry in Japan may not preclude an equally remarkable reconstruction of the domestic of provision in the application of science and technology to social needs. Will visitors from outside, twenty years from now, be beating a path not to the Ministry of International Trade and Industry and its institutions, but to the cradle of what Mr Nakase calls the soft sciences?

The answer depends in part on great imponderables which have puzzled world-wide viewpoints for decades. Why is the Japanese family spirit so easily translated from the home to the office or the factory? Why, in any case, are the Japanese so reliable? But there are also explanations of a more familiar kind, chief among which is that exemplified by the procedures used in the past decades by the Ministry of International Trade and Industry for deciding how to spend public money on technological developments and the machinery then devised for spending the money. One of the most distinguished of all the applied research laboratories in Japan is the Electrotechnical Laboratory in Tokyo, financed by the government and deeply involved not merely in plans for patent recognition, but also in schemes for helping Japanese manufacturing companies to compete still more effectively with manufacturers elsewhere. But the Electrotechnical Laboratory is a pensioner of the Ministry of International Trade and Industry and is one of the agencies by means of which the government can talk to industry about the industrial strategy that Japan should pursue. An enterprise of this kind, if it is a smallish laboratory just under 600 of its members are engaged in research. On page 204, the director of the laboratory, Dr F. Mori, explains how the system works and how he considers that some of its success derives from the Japanese of the Japanese.

But there is a little more to it than that. The Ministry of International Trade and Industry is a powerful government

Japan.

Since 1984, *Nature* has devoted nearly as much editorial effort to Japan as to the United Kingdom, and the journal now carries on average two news articles from Japan every week. The submission of articles and letters from Japanese scientists has also more than doubled in the past three years to more than 500 a year, and hardly a week goes by without one or two research articles from Japan appearing in *Nature*. The circulation is also increasing rapidly.

Nature's editorial efforts in Japan are not just passive. By staging numerous international conferences in Japan, *Nature* has actively sought stronger links with the Japanese scientific community, in particular with young scientists. *Nature's* conferences of 1986, 1988 and 1992 were each attended by about 500 scientists, the majority of them young researchers under 35 years of age supported by industry-sponsored scholarship schemes.

The buffet party of the 1988 conference was graced by the presence of Japan's Emperor and Empress (then Crown Prince and Princess).

It is in this vein that *Nature*, on the basis of its long history of association with Japan, suggests in this special issue what should be the future course of Japanese science as it enters the twenty-first century. □

NATURE • VOL 359 • 15 OCTOBER 1992

JAPANESE SCIENCE POLICY	187
THE SCIENCE COUNCIL OF JAPAN	188
AN OUTLET VIEW	189
UNREST IN THE UNIVERSITIES	193
RESEARCH IN THE UNIVERSITIES	194
NATURAL DISASTERS	197
THE NATIONAL RESEARCH CENTRE FOR DISASTER PREVENTION	200
PREDICTING EARTHQUAKES	202
THE ELECTROTECHNICAL LABORATORY	204
OCEANOGRAPHY IN JAPAN	209
AMINO ACID PRODUCTION BY THE FERMENTATION PROCESS	209
CARDIAC SURGERY IN JAPAN	212
THE PIPS PROJECT	217
JAPANESE MATERIALS ON NEW LINES	218
JAPAN IN SPACE	215
CALCIUM IONS AND MUSCLE CONTRACTION	217
THE NATIONAL INSTITUTE OF GENETICS	219

Many of the concerns expressed in the 1972 special issue remain relevant today.

▷ REFORM

names are listed separately. We are grateful to them for their help. None of them is quoted directly. Their agreement with all or any of the opinions that follow should not be assumed.

Achievements

The implied criticism in what follows does not mean that what has been accomplished by Japan's science in the past two decades is without value. On the contrary, quite remarkable things have been done during those years, as follows:

There is now a network of 14 research institutes, originally intended to provide shared facilities for the universities and a means of educating graduate students, which have become centres of excellence in themselves and which have now spawned their own university (see *Nature* 353, 113; 1991). The electron-positron collider called Tristran at the high-energy physics laboratory KEK at Tsukuba has, for example, made important contributions to the field almost from commissioning day.

Similarly, the Institute of Space and Astronautical Science, originally founded as a part of the University of Tokyo, has retained its reputation for the daring design of interplanetary research experiments. It has become a model of how to carry through low-cost high-benefit space research.

In the same interval, the laboratory called RIKEN, originally founded as a private research institute with the patronage of the Imperial family (see page 578), has grown into a kind of general-purpose basic research organization with no discernible limits in its field of operations, but which is one of several research enterprises in Japan carrying the torch for the internationalization of Japanese research.

Japan has also been extraordinarily active in the creation of new kinds of research support during the past decade. One of the most striking of these is the ERATO project, the brain-child of the Research and Development Corporation of Japan. Groups of up to 20 researchers are recruited from industry, universities and government laboratories to work on deliberately adventurous research projects for five years. Another, is the Strasbourg-based Human Frontier Science Program that supports international research on the brain and biological functions.

Even more spectacular institution-building is represented by the creation of the complex of research institutions at Tsukuba, still sometimes spoken of as Tsukuba Science City, a long hour's bus ride north of Tokyo. Tsukuba has indeed become the technically orientated town-

ship its creators had in mind, but with a sufficient supply of restaurants and other attractions to keep Japan's technocrats happy. It is the chief site for the in-house laboratories of the Ministry of International Trade and Industry, which is now pouring money into their refurbishment. The still-growing university and the scores of industrial laboratories at Tsukuba seem to have grown to the point at which scientists and engineers now change jobs from time to time, leaving one employer for another.

Japan has also gone a long way towards bridging the gap most obvious in the administration of research as recently as 1983 — the lack of postdoctoral positions. The Japan Society for Promotion of Science now provides support for qualified researchers between graduation and their eventual appointment to permanent positions. Not only does this development meet an actual need, but it has also helped to challenge the Japanese tradition that lifelong employment at the same institution is a necessary, perhaps even a sufficient, condition of success.

There is also now in Japan an appreciation that the further improvement of the research enterprise requires further and continuing change. Air travel and language lessons notwithstanding, Japan is still isolated by geography and language. Extra effort is required to be fully a part of the international research enterprise.

New stresses

So things have been moving in the right direction. Why then demand that the process of change should be accelerated? Because the pace of change is not nearly as rapid as it could and should be. And because some of the bastions of tradition and inertia seem unlikely to be touched by a continuation of the reforms of the past two decades.

So much was recognized earlier this year by Japan's Council for Science and Technology, whose Eighteenth Recommendation published in January is entitled "A comprehensive and basic science and technology policy for the new century". The document's most arresting conclusion is that "efforts must be made to double government investment in research and development *as soon as possible*" [emphasis added]. Because the council's chairman is the Prime Minister, Kiichi Miyazawa, and its membership includes four influential members of his cabinet, the promise is seriously regarded; but Japan is still waiting to learn what "as soon as possible" may mean.

In any case, Japan's research (like the rest of Japanese society) is now confronted with unfamiliar challenges (perhaps also opportunities) that will require

the utmost forethought and flexibility in the years ahead. Here are some of the new sources of anxiety:

Recession For the first time in four decades, Japan seems to be heading for an economic recession of unknown severity and duration. Among advanced countries, Japan is comparable only with Switzerland in the degree to which research and development is supported primarily by industry and commerce. What will be the consequences for Japan's science if companies cut back severely on their spending? And what, if the recession is long drawn out and government revenues begin falling, will be the consequences for the ministries' direct support of research?

Reform. Japan's government is now persuaded that something must be done to allow Japan's science to join the modern world. One major driving force for reform is the rapid decline in population reform from a peak this year which is forcing universities to become more competitive to attract students.

One official proudly said the other day that "We are on the move". But there is a serious danger that the bureaucracy will misunderstand what reforms are indeed required. In the administration of the universities and of the research grants which some researchers now enjoy, there may now be more money to spend, but the manner of its spending is so tightly controlled by the bureaucrats that the benefits are to a large degree offset by the way in which researchers' spirits are stifled. Japan is well used to the idea that the funds the ministries dispense should be accounted for down to the last yen, but Japan's practice of this rule is an enemy of both efficiency and daring. If the reform process institutionalizes these practices, it may do as much harm as good.

Basic research. Japanese usually accept as true the unfounded rumour that Japanese companies' successes in the past four decades in fields such as steel, shipbuilding, motor manufacturing, electronics and construction stem from the exploitation of basic research carried out elsewhere. The allegation is unfounded because the public secret of the Japanese success has nothing to do with basic research, whatever its provenance, but hangs on Japan's unique system of interpersonal communication within all organizations. That creates consensus and workers' compliance with their bosses' wishes and provides continual in-service training for all. (That is why the bureaucrats in the government offices sit within earshot of each other.) Superb production engineering also matters. So does market daring.

The rumour has been repeatedly raised in rancorous meetings with the United States over the trade imbalance ▷

▷ REFORM

between the two countries. The Ministry of International Trade and Industry now accepts that Japan must support more basic research, and is pushing its own laboratories in that direction. For different reasons (see below), the companies are following suit. But again there is a danger of misunderstanding. What the new enthusiasts for basic research mean by the term is not always clear. The risk is that they may put money into ill-directed laboratory activity when the effort might be better spent on providing students (graduate and undergraduate) with a better understanding of the principles of science.

Internationalization. Among prosperous countries, Japan stands out for its isolation (geographical and by language) and for its lack of international pretensions.

But that is changing. Japan is already the largest single national source of funds for technical assistance in developing countries. Soon, Japan will have to make a foreign policy for itself, at least on questions such as relations with Russia or with what Russia calls its Far East, not to mention China (which the Emperor of Japan will visit later this month).

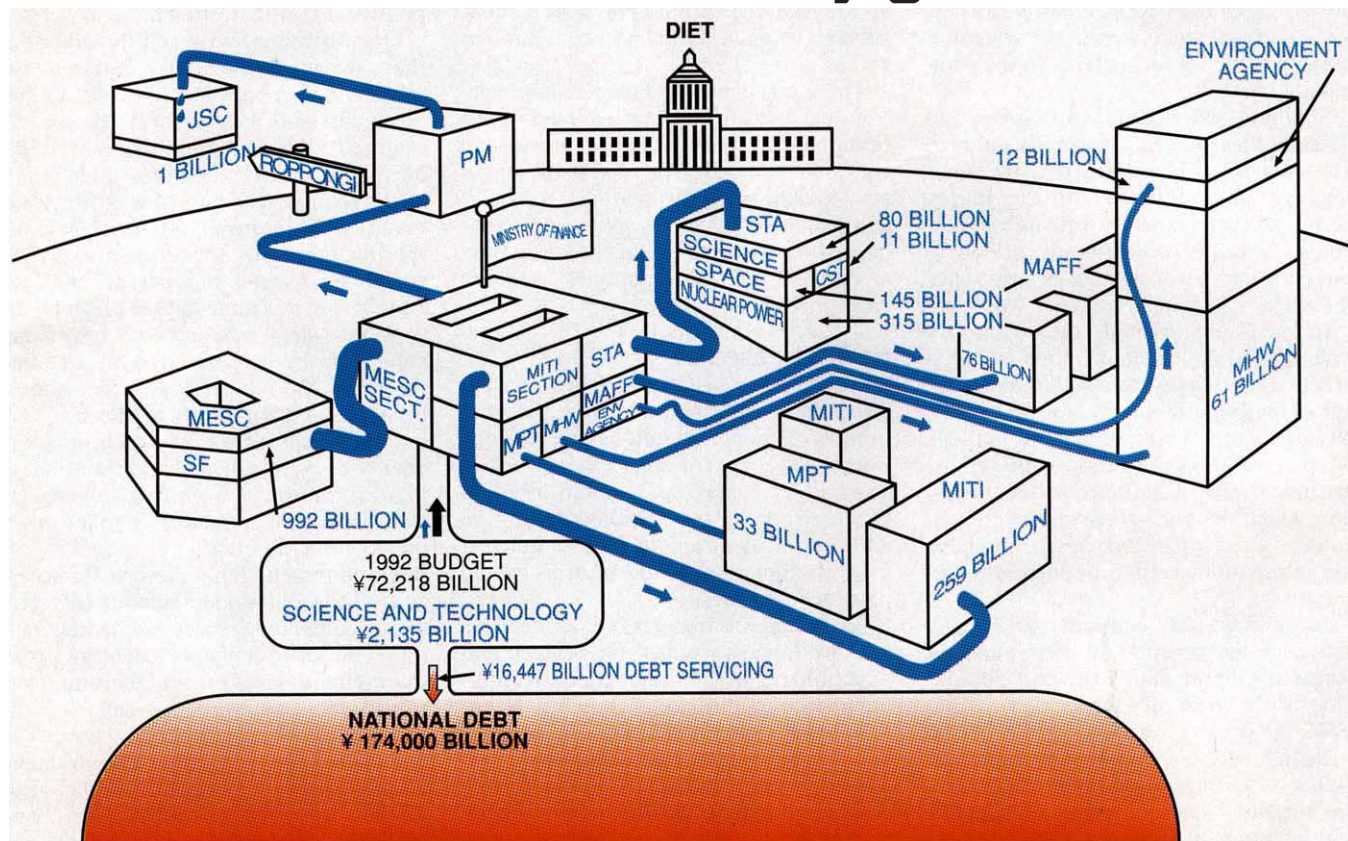
The research enterprise finds itself at the forefront of this movement towards closer involvement with the outside world, and for two reasons. First, the United States has been insisting, at those same rancorous meetings on trade relations, that Japan should "open up" its research laboratories to people from elsewhere. Second, and more telling, reflective researchers know that isolation by language and geography are continuing impediments to membership of the top echelons of the international research

community. That is the spirit in which a seminar the other day at the National Cancer Research Institute (on apoptosis in human glioma cells induced by c-Myc expression) was conducted in English by a Japanese for the benefit of an audience almost exclusively Japanese.

One particularly contentious international issue that will have to be settled by the end of this year is that of whether Japan should contribute towards the cost of the US Superconducting Super-Collider, the particle accelerator now being built in Texas. There are good reasons (see below) why Japan should decline.

Japan's part of the research enterprise cannot but be shaped, in the years and even decades ahead, by powerful trends such as these. How should research policy be devised so as to make the best of them?

Where the money goes



A huge sea of red ink – currently about 174,000 billion yen, (\$1,450 billion) – underlies the finances of the Japanese government and sucks in ever increasing amounts of money from the annual budget to service the growing national debt. As a result, the Finance Ministry has clamped down severely on government spending. Outlays for research have suffered particularly badly over the past decade, and total government spending on science and technology stands at only a little over 2,000 billion yen, or a small fraction of that spent by industry. Most of these limited funds are pumped out of relevant sections of the Finance Ministry to three key recipients: the Ministry of Education, Science and Culture (MESC, or Monbusho), the Science and Technology Agency (STA), within which the Council of Science and Technology (CST) disburses 11 billion yen in grants from STA's 'special Coordination Funds for Promoting Science and Technology' and the Ministry of International Trade and Industry (MITI). Surviving on smaller flows are the Ministry of Agriculture, Forestry and Fisheries (MAFF), the Ministry of Health and Welfare (MHW), and the Ministry of Posts and Telecommunications (MPT). Bringing up the rear is the little Environment Agency, which gets a mere trickle of 12 billion yen a year, while the Science Council of Japan (JSC) in Roppongi struggles to survive on an indirect drip feed of 1 billion yen a year from the Prime Minister's (PM) office. The Cabinet has resolved to double government spending "in the near future". Just when that will be remains uncertain. And whether the present system of pumps and valves for money flow is the most appropriate is also very much open to question.

▷ REFORM

The bureaucracy

The bureaucracy would be comic were it not that it touches every part of Japanese life. Moreover, the term (used as the collective noun for civil servants) is not the term of derision it would be elsewhere; the members of the bureaucracy are proud of their positions, others are appropriately respectful.

Three ministries chiefly affect the administration of science, as follows:

The Ministry of Education, Science and Culture (called Monbusho) administers all education in Japan, spends more than ten times as much on school lunches as on research, finances the 98 national universities (research included) and supervises (and, to some degree, subsidizes) more than 380 private universities.

The Ministry of International Trade and Industry (MITI) does its best to stimulate and coordinate company research (as with the unsuccessful Fifth Generation Computer project), operates a clutch of research laboratories of its own and has a finger in several other pies, the Strasbourg-based Human Frontiers Science Program, for example; and The Science and Technology Agency (STA), whose chief responsibilities are the Japanese programmes in nuclear energy and space, but which spends about a fifth of its budget on good works in general and on Riken in particular.

In addition, the Ministries of Health and Welfare and of Agriculture have quite substantial research programmes, while the Prime Minister's office provides a shoestring budget for the Science Council of Japan, a kind of hybrid of a national academy (whose 210 members are nevertheless elected every four years) and a policy think-tank. (The council's president, Professor Jiro Kondo, is one of the few members of the Prime Minister's Council for Science and Technology.)

The ministries are corporate entities commanding the apparently willing allegiance of their members despite some rigid internal rules (all ministries, for example, rotate their people every two years, with disastrous consequences for continuity). They are rivals in the annual competition for power, influence and budget share, adjudicated by the Ministry of Finance (which second-guesses each ministry with mirroring divisions of its own).

One benefit of the competition is that a ministry may hope to increase its influence by winning funds for new ventures, whence the recent spate of institutional innovation in research. For the same reason, ministries guard with the utmost jealousy what they have already gathered under their control. The top people are enlightened, anxious ▷

Semiprivate research: best of both worlds

Suita. A rich new breed of semiprivate research institutes has emerged in Japan in the past decade which puts Japan's impoverished universities to shame. Among the most lavish are those funded by the Japan Key Technology Center (Japan Key-TEC), an organization set up by the Ministry of International Trade and Industry (MITI) and the Ministry of Posts and Telecommunications (MPT) to divide some of the spoils resulting from privatization of the telecommunications giant Nippon Telegraph and Telephone Company (NTT).

The Protein Engineering Research Institute (PERI) in Suita city in Osaka is MITI's showpiece among the many institutes set up under this organization, while MPT boasts the Advanced Telecommunications Research Institute International, in the fledgling Kansai science city in the hills between Kyoto, Osaka and Nara.

Japan Key-TEC was set up in 1985, with an initial investment of 17,544 million yen (\$140 million) from the government, the Japan Development Bank and private companies to promote collaboration in research between government, industry and universities. Since 1987, MITI and MPT have been pumping 25 to 28 billion yen (\$200–225 million) a year into the organization with funds derived from dividends of government-held shares of NTT.

Japan Key-TEC then channels about 75 per cent of this money into capital investment in research institutes that are jointly set up by two or more companies. The remainder is invested in conditional interest-free loans to companies to support leading-edge research on new technology.

The companies must return the loans with interest if the research produces a successful product, but no interest is payable if the research is commercially unsuccessful. So far Japan Key-TEC has made 230 such loans to companies, for example, to Fuji Heavy Industries to develop robotic aircraft structure assembly, and to Asahi Glass to make chirally pure fluorine-containing amino acids.

Japan Key-TEC's capital investment in institutes, on the other hand, does not have to be paid back. As a result, capital outlays tend to go to more basic research, such as the fundamental work on proteins at PERI.

PERI, with its marble-white walls and an interior decor more reminiscent of a

plush hotel than a research institute, is supported by 14 companies, including Kyowa Hakkō Kogyō, Takeda Chemical Industries, Tonen Corporation, Toray Industries and Mitsubishi Kasei Corporation, with a total fund of 17.1 billion yen (\$137 million) over ten years. Seventy per cent of the budget is provided by Japan Key-TEC and the rest by the companies, who provide researchers to complement those from universities and government laboratories.

The institute, which has about sixty researchers, has the latest NMR equipment and a powerful computer centre, not to mention a surfeit of the most up-to-date equipment for separating and analysing proteins. PERI's ten-year life is supposed to come to an end in March 1996, but Tatsuo Miyazawa, research director of the institute, is confident that the institute can be reborn and continue



PERI — hotel-style surroundings for protein engineers.

under a new name with slightly altered objectives.

The Japan Key-TEC organization, which has set up 83 such semiprivate research laboratories (but none as large as PERI and ATR), is part of an initiative by former prime minister Yasuhiro Nakasone. He set out to develop 'third sector' semiprivate organizations that can tap the rich financial resources of companies and the talents of Japanese academics (who might otherwise be confined to Japan's deteriorating national universities).

The semiprivate laboratories can weather the present economic recession because their budgets are virtually guaranteed over ten years. They also have the flexibility to change direction and re-emerge after their first ten-year life comes to an end. Although it is still a little too early to say how successful they will be, it seems likely that institutes like PERI will assume an increasingly important role in Japan's research and development in the years ahead. □

RIKEN — Japan's leading light

Wako. The Institute of Physical and Chemical Research (RIKEN) in Wako on the northwest outskirts of Tokyo is a rare breed among Japan's government-funded institutes. It enjoys unusual autonomy because it is a 'special corporation' (*tokushuhajin*) funded primarily by the Science and Technology Agency (STA), but it also has other significant independent sources of revenue. The result is a remarkably dynamic research organization that is fast becoming a truly international centre of excellence. As Japanese government officials and academics struggle to create research excellence in Japan's national laboratories and universities, they could do well to learn from the RIKEN model.

RIKEN is not just one research institute. In the past ten years, the Wako-based organization has opened a life science centre in Tsukuba science city northeast of Tokyo and a photodynamics research centre in Sendai several hundred kilometres north of the capital. RIKEN, in collaboration with the Atomic Energy Research Institute, is at present building the world's most powerful synchrotron, Spring-8 (Super Photon ring-8 GeV), in Harima Science Garden City 100 km west of Osaka. Most recently, RIKEN has begun research collaboration on biomimetic control with Nagoya University. RIKEN also has substantial research collaborations with overseas research organizations, such as the Rutherford Appleton Laboratory in the United Kingdom, Massachusetts Institute of Technology in the United States, and the Pasteur Institute in France.

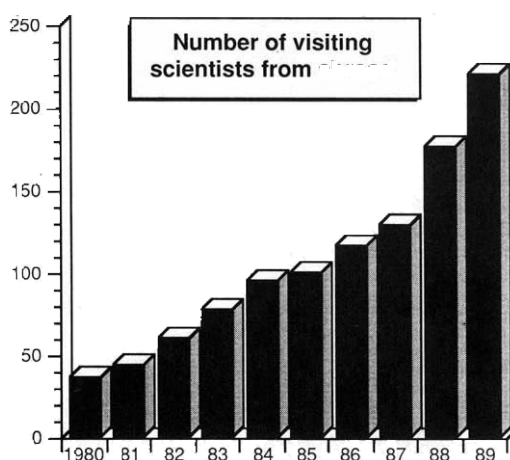
A comparatively new initiative that has undoubtedly added a new international dynamism to RIKEN's activities is the Frontier Research Program. Established in 1986 with funding from STA, the programme runs projects that probe frontier areas of research, such as the brain, new materials and photodynamics. Three key rules make the programme unique: one third of the participants must be non-Japanese, including one third of the project leaders; programme researchers must be young (mainly in their thirties) and are drawn from industry, government and universities; and the projects initially run for only five years, after which they must undergo review and reform if they are to continue.

RIKEN nominally has close to 500 researchers and engineers as permanent staff. But on top of this there is a huge constantly changing population of nearly three times this number of visiting scientists on short-term stays from industry and academic institutions, as well as from overseas. The number of foreign scientists

among them has soared from 37 in 1980 to well over 200 now (see figure), thanks in part to the Frontier Research Program and a new postdoctoral fellowship scheme for foreign scientists introduced by STA in 1988.

Minoru Oda, president of RIKEN, equates his organization to an amoeba that is constantly testing out new ground. Leading the RIKEN organism into new territory is the Frontier Research Program which can put down roots or retract depending on results. But Oda also attributes RIKEN's success to its past history.

RIKEN was established in 1917 as a private foundation with financial support from the Imperial family, government and private organizations. Brilliant scientists, such as Hantaro Nagaoka, whose model of the atom predated Rutherford's, Kohtaro Honda (metallurgy) and Umetaro Suzuki (biochemistry) quickly put RIKEN's name on the world map of science. The institute was visited by Albert Einstein in 1922, and RIKEN has produced two of Japan's five Nobel winners, Hideki Yukawa and Shinichiro



Tomonaga. Patents produced by the institute provided a significant source of operation revenue during its early days.

But RIKEN suffered as a result of the Second World War. The institute was bombed in April 1945. And during the US occupation, Japan's first cyclotron, built by Yoshio Nishina of RIKEN, was ignominiously dumped into the sea by US occupation forces on the grounds that it might be used to develop nuclear weapons. For the next 12 years, the institute went through the worst patch in its history and suffered severe financial difficulties.

However, in 1958, RIKEN experienced perhaps its greatest fortune. The powerful but conservative Ministry of Education, Science and Culture declined

the chance to take control of RIKEN, and the institute was instead put under the wing of the newly fledged STA.

RIKEN now gets a little over 90 per cent of its annual ¥ 19,000 million (\$150 million) budget from the agency. But a significant amount of additional income (more than \$10 million) comes from entrusted research, patent royalties and donations from private industry. This little bit of financial freedom, and STA's wise decision not to pull too hard on the bureaucratic reins of control, has undoubtedly helped RIKEN to flourish.

There are, however, other elements in RIKEN's success. In part because of its impressive pre-war history, the institute has acted as a magnet for attracting leading professors from nearby Tokyo University, Japan's leading national university. Some Tokyo faculties hold visiting appointments. Others have moved permanently to the institute, and yet others are somewhere in between. But, in contrast to Japan's leading national university, all of the principal scientists who head RIKEN's approximately 50 laboratories are subject to a thorough review of their activities every seven years by a committee of scientists from both within and outside the institute. The review committees include members of industry and university researchers, and this review process undoubtedly helps to keep RIKEN's principal scientists on their toes.

RIKEN's researchers have also broken new ground themselves by taking on positions as visiting professors and assistant professors at nearby Saitama University, another first for a government-funded research institute (see *Nature* 351, 175; 1991).

Finally, the recent internationalization of RIKEN must in part be attributed to its current president, Minoru Oda. Before taking up his position at RIKEN, Oda headed the Institute of Space and Astronautical Science, where he initiated several important international collaborations in space science, such as observations of Halley's comet with Japanese, European and Soviet satellites, the Japanese Ginga X-ray satellite that carried detectors from the United States and the United Kingdom, and the Japanese Yokoh satellite that is now carrying out pioneering observations of the Sun's solar maximum with X-ray telescopes developed in the United States and Japan.

Oda is one of Japan's few confirmed internationalists: he knows that the only way forward for his country is through genuine international collaboration on equal terms. □

▷ REFORM

to do the best for their clients; but their pettifogging henchmen are sticklers for the application of the bureaucratic rules.

In relation to science and research, what Japan most needs is an effective coordinating body, able to keep the government as a whole informed of new developments, of individual ministries' responses to them and of matters requiring decision by the government as a whole. But is there not exactly such a body, the Council on Science and Technology (CST), of which no less a person than the prime minister is the chairman?

That is the theory. The practice is that the CST is an awkward hybrid of an advisory council and an executive committee of the Cabinet. The council proper includes not only the Prime Minister but the Minister of Finance and those responsible for MITI, the STA and Monbusho. At intervals of five or seven years the council organizes itself to produce general policy documents, of which the Eighteenth Recommendation published in January is the most recent.

That seems to have been a gigantic undertaking, spread over the eighteen months since the study was commissioned by the then prime minister Toshiki Kaifu, during which six subcommittees held no fewer than 40 meetings of now fewer than 100 members, many of them industrial scientists of distinction. The outcome is a curious mixture of platitude and practicality.

Thus the document defines as one of the major goals for the years ahead the creation of some kind of harmony between science and technology on the one hand, and people and their aspirations on the other, broods on the danger that there may emerge shortages of people interested in careers in science and goes on to urge that "graduate schools must be drastically improved both in quality and quantity". Further exhortations ask for better career structures in science (and equal treatment of women with men) and the famous doubling of the government's own research and development spending, coupled with an improvement of laboratory facilities and of data storage within the publicly supported network.

As a policy document, the Eighteenth Recommendation cannot be faulted for the grandeur of its goals. What minister would decline to put his name to a plan that promises to "settle the global environmental issues, energy issues, food issues, etc., which threaten the existence of human beings . . ."? The deficiencies of the plan in these connections are that it does not explain how these goals might be attained. On the other hand, the

practical recommendations, about budgets and laboratory equipment, for example, remain matters still for decision by the ministries.

It would be better for Japan if the CST were divided into two parts, political and advisory respectively. At the political level, there are legitimate questions to ask (such as whether to join the SSC), while only politicians can decide what can be afforded. But only the technical community can advise on the means by which global ambitions, such as "settling" the world's environmental problems can be achieved.

STA, which houses the office of the Council of Science and Technology, was also supposed to play a coordinating role when established in the late 1950s. But it has become overburdened with development of nuclear power and Japan's space programme, which together take up over 80 per cent of the agency's budget (some Japanese bureaucrats cynically call it the "Space and Nuclear Power Agency"). Given this imbalance and its growing budget, STA's capacity for reflection, mere thinking, on overall government science policy has been diluted.

We conclude that Japan is in urgent need of a more effective coordinating mechanism for science and research. Its objectives would not be the definition and implementation of a single policy, but the exchange of information between sections of the bureaucracy, who are sometimes hardly on speaking terms, mixed with occasional exhortation (on matters such as tenure for foreign scientists). To avoid its being captured by the scent of bureaucratic power, the new body should have no money to spend except the salaries of its small staff.

The universities

Once upon a time, before the Second World War, there were in Japan seven "Imperial" universities, whose statutory function was the training of public servants, and a great diversity of lower-ranking colleges. Although the Imperial universities were regimented places at the outset, by the 1940s, many of them had secured a measure of independence from the centre.

At the end of the war, General Douglas MacArthur and his civil servants declared the system elitist, greatly augmented the number of chartered universities and declared that they should be treated on an equal footing. At the same time, the standard curriculum was transformed: there would be (and, mostly, still are) two years of liberal arts education and two years of major specialization. With two exceptions (Tsukuba University is new and was designed to avoid the problem, and Tokyo University has a strong general education faculty), each of the 98 now 'national' universities is still

seeking its own way to escape this strait-jacket.

Given the post-war terror of elitism, it is curious that the system of national universities that has emerged is one of the most elitist anywhere; all potential students with pretensions to be bright compete for a place at Tokyo.

The most shocking feature (to outsiders) of Japan's system of national universities is the functional separation between academics and their administrator counterparts. Academics appoint their like to permanent positions, and every so often (four years is the standard interval) elect a president, but none of them has the power to appoint an administrator to office.

The senior administrators at the national universities are Monbusho officials seconded for a tour of duty at the coal-face, as it were. One university president claims to be able to veto those thus sent from Tokyo to manage his affairs, but gave no indication of whether he had ever exercised this right. Another said openly that the lack of authority on administration was the most galling feature of his work.

But neither of these presidents nor their academic colleagues can hope to assess carefully the competence of those appointed to do university business at other levels in the academic hierarchy. At the University of Tokyo, the 4,000 tenured academics are almost equally matched by administrative appointments (which, to be fair, include those of nurses at the hospital and workers at the agriculture department's farm).

The second most shocking feature of the national universities is that every academic employee's pay-check is written by Monbusho, not by the university. That practice symbolizes not just the dependence of the national universities on Monbusho, but their national obligations.

The third is the rule of thumb by which the budgets of the universities are determined. Each has a charter, defined by public legislation, specifying the numbers of full professorships or 'chairs' (*koza* in Japanese) in each faculty. The full complement of a chair consists of a full professor, an assistant professor and two assistants.

The rule is that the university is entitled to a certain sum of money each year for each *koza's* running costs. In the science departments of all the national universities, the annual budget in respect of a full professor amounts to about ¥8 million a year, roughly \$65,000. But the cheque is not sent to the professor or even to his department. The university takes the first cut (for library services and for heat and light), the department the next bite and so on. Typically, there may be ¥1 million (about \$8,000) left for ▷

▷ REFORM

a professor and his three colleagues to spend on teaching materials, research expenses and the occasional journey to a conference. Social sciences and the humanities are naturally less well-endowed.

It seems now to be generally agreed that this system is too rigid to make sense. One consequence is inbreeding; an assistant professor worth his salt will expect to fill his retired boss's shoes. Another is fragmentation; on the principle that a boss's inclinations in research should not be inhibited even by his departmental colleagues, strong-minded if eccentric professors have been able to create intradepartmental cells related to the general interests of their departments only by their divergence from them. The system is also, despite the prevailing parsimony, potentially extravagant. The annual cost to Monbusho of a national university is easily calculable from the statutory numbers of *koza* in the various faculties. The total is strictly unrelated to the numbers of students enrolled.

A further structural difficulty is the separation of the four-year undergraduate course into two halves. But this barrier is now being broken down.

There is also the fashionable problem of the 'graduate school'. Many senior academics, with an eye on the research universities of the United States, believe that the national universities need a mechanism for training students in research, without the encumbrance of responsibility for undergraduate teaching.

The University of Tokyo's academic lawyers have devised a neat variant of this case by arranging to have themselves renamed teachers in the Graduate School of Law, whereupon the university has argued that they must also carry out extra duties in the undergraduate faculty; their annual *koza* allowances have already been increased by 25 per cent; those in the science and engineering faculties will increase by the same proportion within the year.

Monbusho, while as vigilant as ever in the counting of the last yen, has been indulgent of this pressure from academics for structural change. Within reason (and the law) it seems now to be saying that universities are free to organize themselves as they choose.

So are the national universities now free? In no sense that would be so understood elsewhere. One university president says that his university's autonomy amounts to the freedom to make academic appointments (subject to the formal approval of Monbusho, never withheld), to change the teaching pattern (subject to the less formal willingness of the academic staff) and individuals' right to pursue what research projects they

choose. That some, perhaps many, pursue no research at all is an open secret.

We conclude that Japan should promptly abandon the convention that all 98 national universities are intrinsically equal, select a small group (say, the seven old Imperial universities plus the Tokyo Institute of Technology) and make them self-governing entities, continue to support them with public funds and hope that from among them will emerge one capable of becoming a commanding centre of international scholarship.

To be fair, there are few in Japan, even among academics, who understand the concept of a self-governing public university, a community of scholars with responsibility both for the education of the young and the management of their own affairs which is supported partly by the government on account of the public service it renders. So familiar is Monbusho's influence that some evidently believe it to be unavoidable and permanent.

There are also tactical reasons why academics at the national universities do not talk much about self-government. "There's nothing Monbusho would like better than to turn us into private universities", said one. "They're much cheaper for the government." But this response is beside the point. There is no reason why a self-governing university should not be supported from public funds on much the same scale as now.

The benefits would seem self-evident elsewhere; self-determination would give academics self-respect, would in principle allow a degree of flexibility in academic planning unobtainable in the present rigid system, would also encourage that self-determination in principle allows and would encourage benefactors from industry and commerce to make donations untied to specific purposes (but for buildings, for example). Given the cosmopolitanism rife at the world's great universities, from Harvard to Heidelberg, it is impossible to see any Japanese university joining that company unless it enjoys a sense of freedom now denied.

Of course, there are also hazards in self-government. A university may make bad judgements, and suffer financially or academically. It may postpone decisions, with similar results. And the receipt of public money entails that the recipients should be accountable — as things are, in Japan, to Monbusho's officials, in a self-governing university, to academic colleagues (which may be an even more painful process). Some are alarmed at the magnitude of these risks. Fortunately, there are some prepared to face them in the hope that some handfuls of excellent researchers will not forever be condemned to work in institutions of studied and contrived mediocrity.

Research

Who can wonder that, in circumstances such as these, academic research in Japan has traditionally been shoestring research?

Although, in the 1960s (the old hands say), a *koza's* stipend could support meaningful research, that is no longer the case. In recent decades, successful research has required support either from a national project of some kind, perhaps based on an on-campus institute, or from an industrial partner interested in employing the department's graduates or from the limited number of large competitive grants.

The charitable explanation of this parlous state of university research in Japan is that Monbusho has had other problems (primary and secondary schools, and their teachers' unions) on its mind. When it might have made a start on university research two decades ago, it was embarrassed by the student rebellion of 1971–72 — the occupation of several departmental buildings at the University of Tokyo for months on end can hardly have made the Ministry of Finance sympathetic towards a plea for research money.

Now there is a plan (see *Nature* 355, 99; 1992) to renew the infrastructure of the universities with the proceeds of selling surplus land that would ordinarily have been returned to the Ministry of Finance, but which Monbusho will be allowed to keep for a provisional period of five years. The University of Tokyo is the first beneficiary.

This supposedly once-only injection of funds (¥100 billion over five years) is only a little more than Monbusho's annual spending on public works at the national universities and laboratories for which it is responsible. It remains a mystery why the bureaucracy that willed the buildings in the first place neglected to provide for their maintenance.

There is also now a research fund, amounting to ¥65 billion this year, and likely to be increased from next April to ¥75 billion a year. Significantly, the fund is called a 'grants in aid'; it supplements Monbusho's payments to universities on behalf of each full professor. Although the sums awarded are not large (typically a few million yen a year) they are usually larger than the sums left over after Monbusho's *koza* payment has been top-sliced and distributed. Everybody agrees that academic research in Japan would be dead in the water if the fund did not exist.

But the supplementary fund is a mixed blessing. There are several types of grants. The supplementary fund thus allows for up to a dozen awards each year to scientists of international standing (worth a few hundred million yen a year for three to five years), 'priority' grants

▷ REFORM

to groups of scientists at different institutions for coordinated work on a common theme (analogous to the similar scheme operated in Germany by the *Deutscher Forschungsgemeinschaft*) and grants for general scientific research awarded to individual researchers.

The importance for the continued improvement of this quickly growing fund is nowhere denied, but it is a bureaucratic nightmare. The general grants, with which ordinary academics are most familiar, are thick with bureaucracy. Many awards are for just a year. Funds may not be available until the summer and must be fully spent by the following February. Spending on technical assistance, one of the most clamant needs, is not allowed, nor on travel for, say, graduate students (even within Japan).

Grant-holders are thus often compelled to cheat; disallowed spending on, say, capital equipment may be hidden by an invoice for computer discs, an enlightened senior person may quarry money for a student's travel by inflating his own travel expenses. Such stratagems may require the connivance of the local bureaucrats, putting the academics concerned at an ethical disadvantage. And most of these grants are very small — typically of the order of ¥1 million a year. They may be a lifeline for academic research in Japan, but they are only a slender lifeline.

We therefore conclude that the administration of these research grants should be separated from Monbusho and put in the hands of a new independent organization able to function as an independent source of research funds — a science foundation or a research council. Among modern industrialized states, Japan is alone in lacking such an organization.

The benefits could be quick. The bureaucracy might be banished. So also might be the present irksome restrictions on how research money may be spent. (The rule that research grants cannot be used to pay for technical assistance stems from the government's overall ceiling on public employees, for example; an independent foundation might be more free than Monbusho.)

But the most important benefit of an independent research-grant organization would be a more discriminatory procedure for awarding grants-in-aid. Some in Japan complain of favouritism in the award process, but the more serious difficulty is that Monbusho, used as it is to dealing with all 98 national universities on an equal footing, cannot also convincingly advocate policies for the concentration of research funds in the universities and departments where they are likely to be used most productively.

The same organization could also

sharpen the rest of the grant-in-aid programme. Thus Monbusho's scheme each year to provide generous research grants, worth perhaps ¥100 million a year, to a dozen or so researchers of international standing would be the more persuasive if the refereeing process involved researchers from outside Japan in a way that an independent agency could manage more easily than Monbusho.

Furthermore, the new organization with a greatly enlarged grant budget could make larger allowances for departmental and university overhead in each grant, thereby giving departments and universities a strong incentive to attract the brightest researchers. Rather than the present 5 per cent overhead, 20 per cent seems more appropriate.

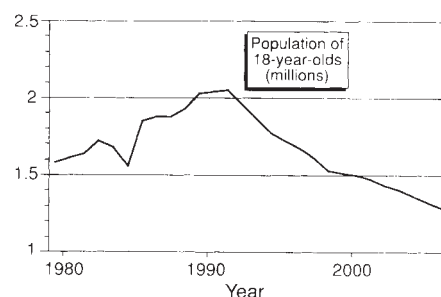
Properly administered (with the help of overseas advisors), the new organization could help gradually weed out poor departments and institutes in the university system by starving them of funds.

Basic research

The CST's Eighteenth Recommendation argues strongly for the reinforcement of Japan's efforts in basic research, and on two grounds. First, Japan has a moral responsibility to contribute to the international stock of knowledge. Second, recognizing that basic research is the fountainhead of innovation in the useful arts, Japan's long-term self-interest requires an indigenous flow of knowledge to found and sustain the industries of the future. How is Japan's research establishment responding? At the policy level, the goal of doubling government spending on research and development is dramatic proof that the case for more basic research is accepted; that will be all the more persuasive when the timescale of that growth has been defined. But it is relevant, and also important, that the message has also apparently impressed itself on industry. The motives are often severely practical. Thus one large electronics company, which operates a successful scheme for recruiting foreign scientists, says frankly that it looks to these recruits for the general understanding of scientific principles in which its indigenously recruited scientists and engineers are often deficient. Despite the remarkable contributions of many companies to basic understanding — Hitachi's contributions to the measurement of microscopic fluctuations of magnetic fields on semiconductor surfaces are but one example — Japanese companies remain convinced that they are better at exploiting new phenomena than at understanding them. That is one reason why NEC (originally the Nippon Electric Company) regards its handsome new laboratory at Tsukuba as Japan's embryonic equivalent of the AT&T Bell Laboratories in the United States. The

score or so of companies that have set up laboratories overseas, mostly in the United States and Britain, are also looking for a harvest of basic research.

The new semiprivate laboratories springing up in Japan, in partnership between industry and government (see page 577), are similarly meant to deepen understanding as well as to offer their member companies routes to practical development of processes and products. Even MITI, whose traditional way of working has been to use small sums of public money to persuade groups of companies to work in concert on common objectives, seems to have embraced the cause of basic research with enthusiasm. One ministry official says that the research programmes of the ministry's



own laboratories are being redesigned around enquiries whose practical benefits cannot be foreseen. Even so, industry is not satisfied. The influential organization of manufacturing industries known as Keidanren has been grumbling for the past year that Japan must strengthen its efforts in basic research, especially at the universities.

The worry is strengthened by the unavoidable decrease of the annual age-group of potential university students (see figure) and by the tendency of young people to opt for studies other than science and engineering. Will there be the young people to sustain Japan's industry a decade or so from now?

Industry's influence on the pattern of research is strong enough to be powerfully influential. Taken together, the direct subvention of university research by Japanese companies is almost comparable with Monbusho's annual grant-in-aid to university researchers. Although companies' support is skewed towards the support of meetings, conferences and overseas travel, much of it is still anchored in the relationships cultivated over the years by companies with individual university departments. In reality, it is improbable that companies would be willing to subsume their special relationships with university departments in a more systematic way of channelling their support to universities.

We conclude that the government, ideally through CST, Monbusho and MITI, should negotiate with indus- ▷

▷ REFORM

trial companies standard terms of research contracts for various purposes (including the payment of overheads to the recipient institutions where appropriate) and that the feasibility of an industrially supported research foundation should be explored.

Internationalization

The Eighteenth Recommendation of the CST on science policy tackles Japan's international connections directly. On the one hand, the document says, there is a need for closer integration with the world's scientific community. On the other, it says, Japan has a special responsibility for using science and technology to improve the lot of developing countries.

For the past decade, the first of these questions has been a constant source of tension, notably because the United States has consistently pressed in international trade negotiations for more open access to Japanese laboratories. The narrow answer to the question is that several government agencies now offer generous postdoctoral fellowships to scientists from overseas. The demand does not meet the supply on offer.

The explanations (language, schooling for young children, jobs for spouses and the like) are well-known. Research organizations such as Riken (see page 578),

but also companies such as Sony, are imaginatively seeking to overcome these cultural difficulties, but there is a long way to go. In any case, the impediments to choosing a long-term career in Japanese research remain substantial.

This is especially true of the universities, of whom only a handful have followed the University of Tokyo in agreeing that overseas scholars should be employed on equal terms with Japanese. Given the polyglot nature of the faculties of the world's outstanding universities, that is short-sighted to say the least. Should not Monbusho take a firm line on this important issue?

The other contentious issue of the immediate present is Japan's membership of the Superconducting Super Collider project (SSC). Even Japan's high-energy physics community is divided on the question (although some government officials represent that it is united behind the project). Most other working scientists are privately suspicious of the project, but unwilling to say so openly out of deference for the still-revered United States.

One objection is that the project is already designed, so that Japan's potentially distinctive contribution to the construction of the accelerator will be negligible. Another is that the management of the project, already in place, does not resemble that of other international accelerator projects (such as those at Gene-

va) in which all partners can expect to have a voice.

But Japanese scientists are also deeply suspicious of the domestic consequences of participation. Playing a full part in SSC would either require an increase of the number of high-energy specialists (estimated at 400 at present) or the postponement of indigenous plans (that for building a prolific source of *B-mesons*, for example). And even if the funds for SSC (estimated at \$2 billion) are taken from some other budget than one used chiefly for science, it is unlikely that the Ministry of Finance would fail to count it in the proposed doubling of the research and development budget.

At a deeper level, it is a matter of pride. Japan, while perhaps over-modest in its estimation of its own basic research, and still admiring of and grateful to the United States, cannot understand why its tax revenues and its skill should be used to fill gaps in a project designed by others without consultation in advance. The chief casualty of compliance could be a loss of Japan's strong trust in the US scientific community.

We therefore conclude that Japan should not join the SSC project, but that if the political pressures at irresistible, the funds should come from some part of the government's budget entirely unrelated to science and research. (The Overseas Development Assistance budget has many advantages.) □

Acknowledgments

We (John Maddox and David Swinbanks) would like to thank the following 65 people who gave freely of their views and time (often at short notice) during the period 10–21 September. We would also like to thank Rinoko Asami who had the difficult task of arranging the schedule and ensuring that we were where we were supposed to be at the right time. Names are listed in the order met.

Katsuhide Kusahara, director of Personnel Division, Ministry of Education, Science and Culture.

Masahiro Kawasaki, senior vice president, Research and Development Corporation of Japan.

Saburo Nagakura, president, Graduate University for Advanced Studies.

Atsuo Nakanishi, secretary-general, Chemical Society of Japan.

Shunzo Ishihara, director-general, Agency of Industrial Science and Technology (AIST), Ministry of International Trade and Industry.

Yoshio Tanzawa, director, Technology Research and Information Division, AIST, Ministry of International Trade and Industry.

Jiro Amagai, director, General Coordination Division, AIST, Ministry of International Trade and Industry.

Osamu Moriya, deputy director, General Coordination Division, AIST, Ministry of International Trade and Industry.

Tsuyoshi Nakai, director for International Cooperation of Technological Development, AIST, Ministry of International Trade and Industry.

Masaya Yasui, deputy director, Technology Policy Research and Analysis Division, AIST, Ministry of International Trade and Industry.

Akira Itoh, director, Superconductivity Division, AIST, Ministry of International Trade and Industry.

Wataru Mori, member of Council of Science and Technology.

Akiyoshi Wada, emeritus professor, Tokyo University.

Ikuo Kushi, dean of Faculty of Science, Tokyo University.

Steve Yamamoto, professor of Kanagawa University, and former professor, Tokyo University.

Muneyuki Date, dean of Faculty of Science, Osaka University.

Kozo Iizuka, executive managing director, Kubota Corporation.

Jiro Kondo, president, Science Council of Japan.

Minoru Oda, president, Institute of Physical and Chemical Research (RIKEN).

Steven Baker, R & D manager, Sony, manager, Coordination Department, Corporate Planning Group.

Andrew House, Corporate Communications, Sony.

Settsuro Ebashi, director-general, National Institute for Physiological Sciences.

Junichi Nishizawa, president, Tohoku University.

Hirohisa Hiramatsu, director, Ceramics and Construction Material Division, Ministry of International Trade and Industry.

Genya Chiba, vice president, Research and Development Corporation of Japan.

Masahiro Itoh, director-general, Frontier Research Program, Institute of Physical and Chemical Research (RIKEN).

Takashi Sugimura, president emeritus, National Cancer Center Research Institute.

Yasutomi Nishizuka, professor, Department of Biochemistry, Kobe University.

Hisatsune Watanabe, general manager, NEC Microelectronics Research Laboratories.

Kosuke Hori, Member of House of Representatives.

Kaname Ikeda, director Personnel Division, Science and Technology Agency.

Tetsuhisa Shirakawa, director, International Affairs Division, Science and Technology Agency.

Hideki Osada, director-general, Science and Technology Policy Bureau, Science and Technology Agency.

Satoru Ohtake, deputy director, International Affairs Division, Science and Technology Agency.

Katsuhiko Umehara, director, Industrial Policy Bureau, Ministry of International Trade and Industry.

Roy Lang, general manager, NEC Fundamental Research Laboratories, Tsukuba.

P. M. Ajayan, NEC Fundamental Research Laboratories, Tsukuba.

Katsumi Tanigaki, research scientist, NEC Fundamental Research Laboratories, Tsukuba.

Thomas Ebbesen, research scientist, NEC Fundamental Research Laboratories, Tsukuba.

Yasuo Hara, vice-president, Academic Affairs, Tsukuba University.

Yasuo Nannichi, vice-president, Research and International Affairs, Tsukuba University.

Motokazu Hori, vice-president, Medical Affairs, Tsukuba University.

Leo Esaki, president, Tsukuba University.

Seigi Iwata, head of physics department, National High Energy Physics Laboratory (KEK), Tsukuba.

Robert Lewis, Tsukuba Research Consortium.

Sonoko Tsukahara, director, ULVAC Tsukuba Laboratory.

Hiroaki Sugino, chief researcher, Stanley Electric Tsukuba Laboratory.

Tsuyoshi Hayakawa, Hamamatsu Photonics, Tsukuba Laboratory.

Eiichi Ozawa, research manager, Air Liquide Laboratories, Tsukuba.

Tetsuzo Kawamoto, managing director, Tsukuba Research Consortium.

Susumu Nishimura, executive director, Banyu Tsukuba Research Institute.

Akito Arima, president, Tokyo University.

Robert Geller, associate professor, Earth and Planetary Physics, Tokyo University.

Tatsuo Miyazawa, research director, Protein Engineering Research Institute, Osaka.

Haruki Nakamura, director, second research department, Protein Engineering Research Institute.

Nobuhiro Go, professor, Department of Chemistry, Kyoto University.

Junjiro Kanamori, president, Osaka University.

Kenichi Matsubara, Institute for Cellular and Molecular Biology, Osaka University.

Yoshiyuki Matsuo, editorial staff, Nikkei Science.

Chikara Hayashi, chairman, ULVAC.

Teruo Hiruma, president, Hamamatsu Photonics.

Kazuo Wada, general manager, president's office, Hamamatsu Photonics.

Yukio Sugino, member of the board, Pharmaceutical Group, Takeda Chemical Industries.

Yoshio Nakao, managing director, Senri Life Science Foundation, Osaka.

Shozo Ino, professor, Department of Physics, Faculty of Science, Tokyo University.

Messing around with gravity

Attempts to dispense with the need for galactic dark matter by altering the inverse square law of gravitational attraction have attracted a small but persistent advocacy.

THE simplicity of classical gravity, whether formulated according to Newton or to Einstein, is both pleasing and puzzling. It is pleasing because we want our physical theories to look good, but puzzling because we always then ask about the origin of the very simplicity that pleases us. To the questions why gravity should exist at all and why it should obey an inverse square law, Newton famously answered *hypotheses non fingo*, meaning (in the context) "Don't ask me". But this is plainly insufficient, at least for those physicists who spend their time trying out modifications to classical gravitational theories.

Newton's law can be amended (as in the case of the now defunct fifth force) to include some dependence on composition as well as mass, and small departures from the inverse square law are allowed on both the short and the long range. To Einstein's theory, which is the simplest nontrivial theory linking space-time curvature to the mass it contains, can be added not only the well-known cosmological constant but also terms containing higher derivatives of the curvature tensor. Physicists have made all these changes and more, and for more substantial reasons than mere curiosity.

One undisputed body of evidence that may indicate some need for a change to the laws of gravity comes from the measurement of galactic rotation curves. Radioastronomers can detect, beyond the visible limit of galaxies, faint hydrogen emission. By measuring the received wavelength of the Doppler-shifted 21-cm emission line, they can then plot the relative rate of rotation against distance from the galactic centre. From this simple exercise, first performed systematically by Vera Rubin and her colleagues in the 1970s, arises the odd result that galactic rotation tends to level off at some hundreds of kilometres per second — and to stay there.

Kepler's law says that the rotation velocity beyond the limits of a finite body of mass must fall off with distance, and conventional interpretation of a flat rotation curve is that there must be extra mass — dark matter — well beyond the point at which no light is seen from the galaxy. (This is the best evidence for dark matter in the Universe, but for all galaxies it adds up to substantially less mass than would constitute a critical cosmological density; whether there is enough dark matter to provide for that is another question.)

A priori, as the lawyers would say, one could equally well argue that if there is a conflict between one observation and another,

the theoretical connection between them must be suspect. In this case, the flat rotation curve and the apparent absence of matter cause a problem only because of Kepler's law, which derives immediately from Newton's law of gravity. So why not change the theoretical law to leave the observations with their face-value interpretation? There is then no dark matter, but something must be done to the inverse square law.

For years, Mordehai Milgrom of the Weizmann Institute has been trying to make precisely this point. He has proposed changing Newton's law by adding to the usual definition of the force of gravity an extra term depending on the relative acceleration of the interacting masses. No reason is given for making this change, except to explain the rotation curves, but then Newton felt no need to justify his law beyond the fact that it explained the orbits of the Moon and planets — the Solar System's rotation curve, as it were.

Milgrom's theory asks for the gravitational force law to be slightly different when the acceleration is small, which then allows a different rotation curve to be obtained for the stately motion of gas a long way from the centres of galaxies. The occurrence of a term dependent on the acceleration is odd, but perhaps not so odd as it seems, because accelerations have absolute meaning (unlike velocities). On the other hand, Milgrom would have an abrupt change from normal to acceleration-dependent gravity at some suitably adjusted critical value of the acceleration, which makes his theory look untidy.

A more palatable modification of the inverse square law is to add a term reciprocal in distance (rather than the distance squared) that takes over only at large distances. Richard Liboff now explains in *Astrophysical Journal Letters* (397, L71; 1 October 1992) that even this is a little too simple, and cannot explain the whole range of galactic dynamical data. So Liboff attempts instead to make Milgrom's law tidier by including in the force law an acceleration term whose magnitude varies smoothly with distance. The particular mathematical form is entirely cooked up, following from no fundamental principle, but that is the spirit of this game.

Liboff's conclusion reinforces what Milgrom has been saying all along, that rotation curves can be explained with a modified gravity law, but goes beyond it in arguing that certain features of internal galactic dynamics, which gave Milgrom trouble, also fit with the generalized acceleration-dependent law. So why have physicists and cosmologists been ignoring Milgrom all these

years, as they will undoubtedly now ignore Liboff? If the purpose is to explain galactic rotation curves, and if there is no other observational or experimental evidence bearing on the question, why is it acceptable to propose invisible, undetected and unidentified sorts of dark matter but, apparently, not quite seemly to fiddle with the inverse square law of gravity?

Part of the answer is that, in the modern idiom, the interactions between particles are perceived as somehow more elementary than the particles themselves. There are only four forces known to physics (three, if the weak and electromagnetic interactions are counted as having been successfully combined), but there are lots of apparently elementary particles. Theories that unify the forces further, or explain certain puzzling aspects of their symmetries and symmetry breakings, tend to create neater mathematical structures at the expense of adding to the storehouse of particles (Higgs bosons and axions come about in this way). That in turn suggests that it is safer to make up a new particle (dark matter) than to disfigure the mathematical form of a well-known force (as Milgrom's theory does).

Aesthetics thus account for some of the studied neglect of Milgrom's numerous papers, but there is also a more quantitative physical objection to his idea. Newton's law of gravity is successfully subsumed into Einstein's general theory of relativity, and a better understanding of the nature of gravity thereby provided. But nobody has yet found a modified form of relativity that incorporates Milgrom's theory; it remains an orphan in the classical world. And although there have been modifications of the simplest forms of Einstein's equations, none of them has so far proved relevant to the problem of galactic rotation curves. At the same time, it is hard to avoid the conclusion that any theory that leaves the workings of gravity untouched on small scales but makes changes on the scale of galaxies might be expected to imply more substantial changes yet on the cosmological scale.

The fact that Milgrom's theory of gravity has no relativistic counterpart is both an aesthetic flaw and a practical obstacle. It is more than slightly inconsistent to use general relativity to explain the Universe as a whole but a classical law of gravity that does not emerge from relativity to account for galactic dynamics. Until Milgrom has his own version of general relativity, his explanation of rotation curves must seem *ad hoc*. But then nobody has found any dark matter yet either.

David Lindley

Collision complexes clocked in real time

Roger Grice

THE characterization of long-lived collision complexes has long been central to the study of chemical reaction dynamics. Chemists would like to understand the dynamical structure of these highly vibrationally and rotationally excited intermediates in a manner analogous to the determination of molecular structure by chemical spectroscopy. A significant stride in this endeavour has now been taken by Zewail and co-workers¹, based at the California Institute of Technology, who have directly measured the lifetime distribution of collision intermediates for the first time using femtosecond laser spectroscopy.

The idea that highly vibrationally excited intermediates with sufficient energy to undergo dissociation might still persist for many vibrational periods was first proposed in the Lindemann theory of unimolecular reaction kinetics². This concept found confirmation in the first observation of a long-lived collision complex in the molecular beam studies of alkali-atom alkali-halide exchange reactions by Miller, Saffron and Herschbach³. In each of these cases the lifetime of the persistent reaction intermediate was inferred indirectly, the clock being the frequency of deactivating collisions by other molecules in the case of unimolecular reactions and the rotation

period of the collision complex in the case of the measurements performed in reactive scattering experiments.

The development of femtosecond laser spectroscopy and its application to isolated molecules in pulsed supersonic beam expansions by Zewail⁴ offered the prospect of real time measurements of the lifetimes of reaction collision complexes. This was first attempted⁵ by using the photoinitiated decomposition of the $\text{IH}\cdot\text{CO}_2$ van der Waals molecule (see Fig. 1a). Photodissociation of the HI molecule initiated the reaction



and time-resolved laser-induced fluorescence measurements of the appearance of OH products gave estimates of less than 5 picoseconds for the lifetime of the HOCO collision complex. However the configuration of the $\text{IH}\cdot\text{CO}_2$ van der Waals molecule leaves the ponderous I atom in close proximity to the vibrationally excited HOCO complex. It may therefore be subject to vibrational energy transfer in secondary collisions resulting in drastic changes in the collision lifetime.

Conscious of these shortcomings, Zewail and co-workers¹ have now applied an ingenious twist in their study of the reaction



using the photoinitiated decomposition of the $\text{HBr}\cdot\text{I}_2$ van der Waals molecule (see Fig. 1b). Photodissociation of the HBr molecule now allows escape of the light hydrogen atom before the reaction is initiated by collision of the ponderous Br atom with the I_2 molecule. Consequently, the $\text{Br} + \text{I}_2$ reaction occurs as an isolated bimolecular collision, and time-resolved photoionization measurements of IBr products give true estimates (to about 0.5 picoseconds) of the lifetime of the isolated BrII collision complex with a precisely defined internal energy. Indeed, the fact that most of the recoil energy of the initial HBr

photodissociation is carried off by the light H atom gives a rather low collision energy (2.5 kJ mol^{-1} , or so) to the subsequent $\text{Br} + \text{I}_2$ reaction. The observed dependence of the lifetime of the BrII collision complex on collision energy is therefore expected to be very sensitive to the form of the reaction potential energy surface.

Zewail *et al.*¹ use classical calculations of the reaction trajectories to investigate the form of the potential energy surface. They conclude that the reactants may enter the potential energy well which holds the BrII collision complex without any intervening barrier in the entrance channel, but that dissociation of the complex to reaction products is impeded

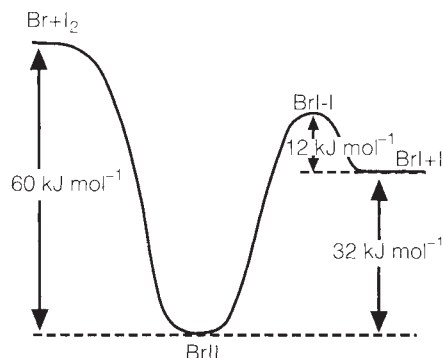


FIG. 2 Potential energy profile for the reaction



showing the potential energy well corresponding to the BrII complex and the potential energy barrier to dissociation.

by a substantial barrier in the exit channel of the potential energy surface (see Fig. 2).

Crossed molecular beam reactive scattering measurements at a higher collision energy (about 12 kJ mol^{-1}) by Lee, Herschbach and colleagues⁶ provided evidence of a short-lived complex with a lifetime of some 5 picoseconds, comparable to the rotational period of the complex. It will be of great interest to see whether this difference in collision lifetime can be attributed purely to the increase in collision energy, because the reactive collisions which give rise to scattering in the cross-beam experiments may extend out to much larger impact parameters ($b_{\text{max}} \approx 4 \text{ \AA}$) than those which

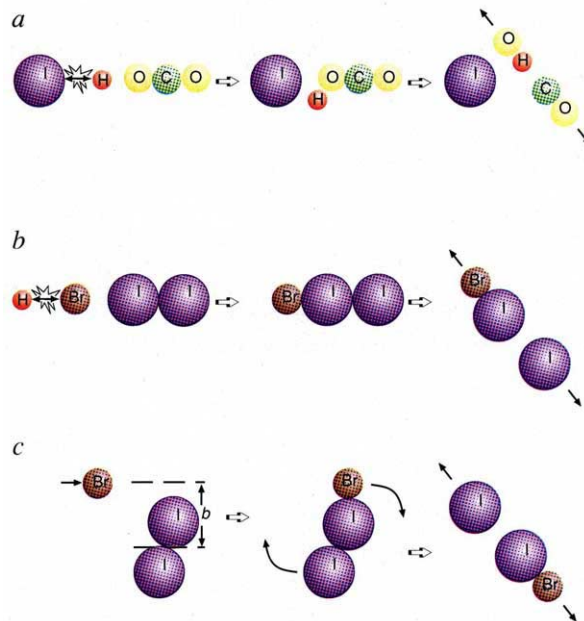


FIG. 1 Reaction configurations for the photodissociation-initiated reactions a, $\text{H} + \text{CO}_2 \rightarrow \text{HOCO} \rightarrow \text{HO} + \text{CO}$ and b, $\text{Br} + \text{I}_2 \rightarrow \text{BrII} \rightarrow \text{BrI} + \text{I}$; and the cross-beam reaction c, $\text{Br} + \text{I}_2 \rightarrow \text{BrII} \rightarrow \text{BrI} + \text{I}$. The impact parameter, b , is the distance between the reagents' centres of mass.

1. Sims, I. R., Gruebele, M., Potter, E. D. & Zewail, A. H. *J. chem. Phys.* **97**, 4127-4148 (1992).
2. Lindemann, F. *Trans. Faraday Soc.* **17**, 598-599 (1922).
3. Miller, W. B., Saffron, S. A. & Herschbach, D. R. *Discuss. Faraday Soc.* **44**, 108-122 (1967).
4. Zewail, A. H. *Science* **242**, 1645-1653 (1988).
5. Scherer, N. F., Sipes, C., Bernstein, P. B. & Zewail, A. H. *J. chem. Phys.* **92**, 5239-5259 (1990).
6. Lee, Y. T., McDonald, J. D., LeBreton, P. R. & Herschbach, D. R. *J. chem. Phys.* **49**, 2447-2448 (1968).
7. Jarvis, R. D. & Grice, R. *Molec. Phys.* **65**, 1205-1216 (1988).
8. Smith, D. J. & Grice, R. *Molec. Phys.* **73**, 1371-1382 (1991).
9. Chatfield, D. C., Friedman, R. S., Schwenke, D. W. & Truhlar, D. G. *J. phys. Chem.* **96**, 2414-2421 (1992).

arise from the constrained geometry of the photoinitiated reaction of the $\text{HBr} \cdot \text{I}_2$ van der Waals molecule (see Fig. 1c). The very high angular momentum ($L_m \approx 250h/2\pi$) of complexes formed in the cross-beam reaction may also exert a strong influence on the lifetime. Indeed, much progress^{7,8} has been achieved in recent years in identifying the structure and dynamics of persistent complexes, and characterizing the transition states for their dissociation to reaction products in reactive scattering experiments.

Zewail and co-workers now hold out the prospect of gaining higher collision

energies by photodissociating the $\text{CH}_3\text{Br} \cdot \text{I}_2$ van der Waals molecule and observing the BrII complex directly, thereby gaining sufficient resolution to identify the time evolution of individual vibration-rotation states of the complex. With these and parallel theoretical advances on quantized transition states⁹, we may truly be within sight of a spectroscopy of reaction transition states and collision complexes. □

Roger Grice is in the Department of Chemistry, University of Manchester, Manchester M13 9PL, UK.

ASTROPHYSICS

Stellar opacities in a flash

D. B. Guenther

THE opacity of matter has been measured for the first time at the temperatures and densities that occur in the interiors of stars. Da Silva, of the University of California at Berkeley, and his collaborators at the Lawrence Livermore National Laboratory¹ used an X-ray laser to heat a thin sample of iron to about 2.5×10^5 K, while X-rays from another laser were used to determine the photo-absorption properties of the heated sample. The entire measurement lasted less than a nanosecond because of the speed of cooling.

Because the densities and temperatures of matter in stellar interiors are inaccessible to experiment (the Sun's central density is greater than 100 g cm^{-3} and its central temperature is greater than 1.5×10^7 K), the photo-absorption properties of stellar matter must be calculated theoretically. For more than two decades, tables of opacities calculated at Los Alamos National Laboratory^{2,3} (LANL) have been used to construct stellar models, which depend enormously on how photons generated in the core find their way out through the highly ionized medium.

At the time that the LANL opacity programs were being developed, computers were not sufficiently powerful to carry out a full and detailed solution to the problem, and several approximations to the atomic physics were made. But armed with the latest supercomputers, Rogers and Iglesias, at the Lawrence Livermore Laboratory, have now calculated a new set of opacity tables (OPAL), that are based on better atomic physics and that take into account the Coulomb interactions among all the electrons and the nuclei of the mixture of elements found in stars⁴.

To the surprise and delight of many astrophysicists, the OPAL opacities are larger than the LANL opacities (up to

three times larger at temperatures near 3×10^5 K). The astrophysicists were surprised because they had believed the LANL opacities to be accurate to within 25 per cent; they were delighted because the increased opacity was just what they needed to resolve several difficulties in stellar-evolution theory.

The list of puzzles that have been resolved is quickly growing, with the most dramatic improvements being found in solar and stellar pulsation theory. One of the first published success stories for the OPAL opacities concerned classical Cepheid and RR Lyrae stars — stars whose brightness pulsates every few days. (Their opaque atmospheres physically inflate as they are heated by radiation from below, until they become transparent enough to let the radiation escape, subsequently deflating, and so on.) Before the OPAL opacities were available, the mass of the stars could be adjusted in astrophysical models to give the correct mean luminosity or to give the correct pulsation period, but not both. With the OPAL opacities, the mass discrepancy disappears⁵. Other success stories include bringing the theoretically predicted non-radial acoustic oscillation spectrum of the Sun into closer agreement with that observed⁶, and removing the need to invoke large amounts of convective-core overshoot in explaining the observed mass-luminosity relationship in massive stars⁷.

An important new feature of the OPAL opacity calculation is the inclusion of so-called $\Delta n = 0$ transitions. These transitions (in which an atom's principal quantum number n remains unchanged, while its angular momentum does change) add a dense forest of lines to the spectrum, especially from iron, which increases the opacity inside the Sun's convective envelope at tempera-

tures near 3×10^5 K. This increase is responsible for most of the improvements in stellar pulsation theory. The critical test for the new OPAL opacities was Da Silva and colleagues' experiment, because it measures the opacity of iron near these temperatures.

To do the experiments required the Nova laser facility at Lawrence Livermore, one of the most powerful lasers in



Roger Ressmeyer, Starlight/Science Photo Library

Star turn: Nova, one of the world's most powerful lasers, used by Da Silva and colleagues in their iron-plasma experiments, is seen here imploding a hydrogen pellet in an inertial fusion experiment.

the world. The 20-nanometre layer of iron (sandwiched in polypropylene supports) was heated by X-rays from a gold film irradiated by a 1-nanosecond, 3.3-kilojoule pulse of laser radiation. Its temperature raised to 2.5×10^5 K, the film's opacity was probed by a second pulse of X-rays, made in the same way, which passed through the plasma and were analysed with a spectrometer.

The agreement between the measured opacity and that predicted is remarkable, and confirms astrophysicists' faith in the OPAL calculations. □

D. B. Guenther is at the Center for Solar and Space Research, Yale University, PO Box 6666, New Haven, Connecticut 06511, USA.

1. Da Silva, L. B. *et al. Phys. Rev. Lett.* **69**, 438–441 (1992).
2. Cox, A. N. & Steward, J. *Astrophys. J. Suppl.* **31**, 243–259 (1970).
3. Heubner, W. F., Merts, A. L., Magee, N. H. & Argo, M. F. *Los Alamos Scientific Report LA-6760-M* (1977).
4. Rogers, F. J. & Iglesias, C. A. *Astrophys. J. Suppl.* **79**, 507–568 (1992).
5. Cox, A. N. *Astrophys. J.* **381**, L71–L74 (1991).
6. Guenther, D. B., Demarque, P., Kim, Y.-C. & Pinsonneault, M. H. *Astrophys. J.* **387**, 372–393 (1992).
7. Stothers, R. B. & Chin, C.-W. *Astrophys. J.* **381**, L67–L70 (1991).

No muscles, but what a brain

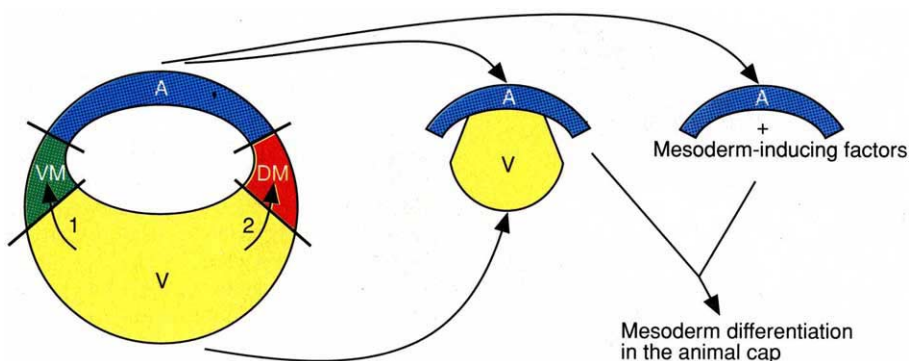
Patrick Lemaire

THE article by Hemmati-Brivanlou and Melton found on page 609 of this issue¹ dissipates growing doubts about the role of activins in the formation of mesoderm in the equatorial region of early *Xenopus* embryos. By showing that a truncated activin receptor can act as a dominant negative mutant and block mesoderm formation, Hemmati-Brivanlou and Melton demonstrate that molecules of the activin family do not merely mimic the endogenous mesodermal inducer but are in fact some of its necessary components. This in itself would be a fine and extremely valuable result. Yet a further analysis of the phenotype of embryos injected with the dominant negative receptor is a source of considerable surprises that may eventually lead to a new understanding of the formation of mesoderm and neural tissue.

The study of mesoderm formation in the amphibian *Xenopus laevis* provides an excellent model system for understanding the events leading to the establishment of the vertebrate body plan. Before mesoderm induction a *Xenopus* blastula is composed of two cell types: presumptive ectoderm in the animal hemisphere and presumptive endoderm in the vegetal hemisphere. Mesoderm arises from animal cells in the equatorial region as a result of their induction by vegetal cells. Dorsal vegetal tissue induces dorsal mesoderm (and the Spemann organizer), whereas ventral vegetal tissue induces ventral mesoderm² (see figure).

A milestone in the identification of the signals emitted by the endoderm was the discovery³ that mesoderm was induced in animal caps treated with a supernatant from a *Xenopus* cell line. The active molecule was subsequently shown to be an activin⁴, a polypeptide growth factor of the transforming growth factor- β (TGF- β) family, that can also induce the release of follicle-stimulating hormone and promote erythroid differentiation. Other inducing factors were identified, all belonging to the TGF- β or fibroblast growth factor (FGF) families (for review see ref. 5), and the FGF pathway was shown to be necessary for ventral and posterior mesoderm formation in the embryo⁶. All mesoderm-inducing factors can induce ventral mesoderm in animal caps, but activins are the only molecules able to induce dorso-anterior tissues such as notochord. An activin-like activity was detected in early embryos⁷, making activins the only obvious candidates for the dorso-anterior mesoderm inducer and hence generating a great deal of enthusiasm.

Unfortunately, doubts soon started to accumulate. If activins provide the dorso-anterior signal, their overexpression in the ventral vegetal side of an embryo should lead to the induction of an ectopic Spemann organizer and the subsequent development of a complete secondary axis. In fact, it was found⁸ that the secondary axis induced by activin lacks anterior structures. Therefore the dorsal inducer cannot consist of activins alone.



Cross section through a *Xenopus* blastula, illustrating emission by the vegetal hemisphere (V) of two qualitatively or quantitatively different signals. Signal 1 will induce ventral mesoderm (VM) whereas signal 2 will induce dorsal mesoderm (DM). This induction process can be reconstituted *in vitro* by dissecting the animal cap (A) and putting it in contact with a piece of vegetal tissue or treating it with mesoderm-inducing factors such as activin or fibroblast growth factors.

They might even be unnecessary, in that the recent identification of two additional factors suggested an alternative route to dorsal mesoderm. Molecules of the Wnt family and a new polypeptide named Noggin have the striking ability to induce a Spemann organizer when overexpressed in ventral endoderm^{8,9}. They have no apparent mesoderm-inducing activity but can collaborate with FGF to generate dorsal mesoderm^{9,10}. In the embryo a combination of a ventral mesoderm inducer with one of these factors could replace activins.

Thanks to Hemmati-Brivanlou and Melton¹, these agonizing doubts about activins belong to history. The clear answer is that a ligand for an activin receptor is needed for mesoderm induction. It might of course be argued that the receptor could bind a ligand of a different family, but this seems to be unlikely. Now that the need for an activin-like molecule in mesoderm induction during normal embryogenesis is established, the hunt for the molecule will be on and is likely to meet with success. But even before the nature of the ligand(s) is known, the new results provide us with insight into their role in early development. Here the surprises begin.

As the FGF pathway is not disrupted in embryos overexpressing the dominant negative mutant, the endogenous FGF would be expected to induce ventral or posterior mesoderm in the equatorial region of these embryos. All the injected embryos should therefore contain some ventral or posterior mesoderm. In fact, many of them do not contain any mesoderm detectable by histology and do not express Brachyury, a general mesodermal marker. This indicates that whereas FGF alone is sufficient to induce mesoderm in animal cap cells, it needs to collaborate with activins to do so in cells fated to become mesoderm in the equatorial region of the embryo.

To me, these results are a sign that our understanding of mesoderm induction has progressed to a stage where we begin to see the limitations of the useful and convenient animal cap assay, and can no longer consider that it accurately reflects what is happening in the equatorial region of the embryo. To understand the interactions between the pathways of the different endogenous mesoderm inducers, we will have to switch to the difficult study of normal presumptive mesoderm cells.

Questions also remain concerning the ontogeny of dorso-anterior structures. Do they arise as the result of the collaboration of activin with a molecule with properties similar to Wnt or Noggin, or with an (as yet unidentified) endogenous dorsal mesoderm inducer? The answers to these questions probably lie in a combination of approaches involving different organisms. The simplicity of the organization of the early *Xenopus* embryo, its large size and the availability of new lineage-tracing techniques¹¹ will favour biochemical and embryological approaches. But it is the generation of mutant mice that lack potential endogenous factors or their receptors that should provide conclusive evidence of the involvement of individual molecules.

The second unexpected result is that the activin pathway could also be important in neural induction. The current view of neural induction is that ectoderm cells need to receive a signal, vertical or planar, from mesoderm to differentiate into neuroectoderm. But animal caps isolated from embryos injected with the dominant negative activin receptor spontaneously express neural markers such as the neural cell adhesion molecule N-CAM without any signal from the mesoderm. Surprising as it may seem, this result can be reconciled with the present model if it is postulated that ectoderm cells are kept along an epidermal differentiation pathway by an activin-like molecule. The neural inducing signal could then be an antagonist of activin action and the authors mention two such compounds, follistatin and inhibin, as good candidates.

I am already looking forward to reading about the effect of these two factors, but I also wonder what could be the nature of the activin-like substance that keeps animal cap cells in an epidermal state. Is it the same one inducing the

mesoderm (then, if so, why should it be present in the whole animal cap)? Or is it different? However that may be, the report by Hemmati-Brivanlou and Melton, by raising many provocative issues and by offering hints of ways to tackle them, will cause a brainstorm among those working on early inductive events.

Patrick Lemaire is in the Wellcome/CRC Institute of Cancer and Developmental Biology, Tennis Court Road, Cambridge CB2 1QR, UK.

1. Hemmati-Brivanlou, A. & Melton, D. A. *Nature* **359**, 609–614 (1992).
2. Dale, L. & Slack, J. M. W. *Development* **100**, 279–295 (1987).
3. Smith, J. C. *Development* **899**, 3–14 (1987).
4. Smith, J. C., Price, B. M. J., Van Nimmen, K. & Huylebroeck, D. *Nature* **345**, 729–731 (1990).
5. Jessel, T. M. & Melton, D. A. *Cell* **68**, 257–270 (1992).
6. Amaya, E., Musci, T. J. & Kirschner, M. W. *Cell* **66**, 257–270 (1991).
7. Asashima, M. *et al. Proc. natn. Acad. Sci. U.S.A.* **88**, 6511–6514 (1991).
8. Sokol, S., Christian, J. L., Moon, R. T. & Melton, D. A. *Cell* **67**, 741–752 (1991).
9. Smith, W. C. & Harland, R. M. *Cell* **70**, 829–840 (1992).
10. Christian, J. L., Olson, D. L. & Moon, R. T. *EMBO J.* **11**, 33–41 (1992).
11. Nierhs, C. & De Robertis, E. M. *EMBO J.* **10**, 3621–3629 (1991).

ATMOSPHERIC EVOLUTION

Life and the rock cycle

Ján Veizer

UNRAVELLING the origin and history of life within the context of planetary evolution is a formidable challenge. The discovery of life forms in rocks 3.5 billion years old, combined with isotope evidence, attests to the existence and ubiquity of life almost as far back as the surviving rock record^{1,2}. That this life interacted with air, water and rocks — the 'Earth system' in the language of global change — is accepted as self-evident, albeit with a proviso that the inorganic world represents only a stable substrate for the dynamic and rapidly evolving biosphere, the latter even shaping the Earth system for its own benefit³. Yet, as discussed by Des Marais *et al.* on page 605 of this issue⁴, it may in fact be the biochemical basis of life that is unchanging or stable, striving to defend its integrity in the face of protracted pressures imposed on the Earth system by geological evolution.

The consensus two decades ago, influenced by the work of S. Miller and H. Urey, assumed that the 4.5-billion-year-old Earth had a strongly reducing atmosphere that enabled the rise of life sometime in the Archaean (over 2.5 billion years ago). This biosphere gradually expanded and produced biological innovations, such as photosynthetic splitting of the CO₂ molecule that resulted in oxygenation of the atmosphere, precipitation

of banded iron (oxide) ore formations and a host of related phenomena⁵. Subsequent field work^{1,2} unearthed stromatolites (fossilized algal mats) and still-disputed fossil blue-green algae, and pushed back the origin of life a billion years, to at least 3.5 billion years ago. The physical evidence, however, did not yield any information as to the size (mass) of this ancient biosphere, although the comparable organic-carbon contents in Archaean and Phanerozoic sediments hint that it might have been substantial.

In 1970, W. Broecker⁶ pointed out that a simple mass-balance calculation for carbon isotopes can answer such a question, because photosynthetic biological pathways preferentially use ¹²C, leaving the residual ocean-atmosphere system enriched in ¹³C. This way of reasoning was extended by C. Junge and by M. Schidlowski and associates⁷ to show that — almost as far back as we have a geological record — carbonates (as a proxy for past oceans) and organic matter have carbon isotope signatures comparable to their modern counterparts. To be sure, such calculations take into account only the relative sizes of the oxidized and reduced (equals organic) reservoirs of carbon, and the problem is that the actual extant biomass, as opposed to its 'dead' counterpart in

sediments, is a minuscule fraction of the organic reservoir. Furthermore, the scatter in the organic carbon-isotope signal is considerable. Nonetheless, the overall impression is that the primitive ancient biosphere may have been of near-modern size.

If ancient biochemistry operated in a near-modern mode, how can this be reconciled with the geological evidence of a lower partial pressure of oxygen in the ocean-atmosphere system until around 1.8 billion years ago? The discovery of 'black smokers' in the deep sea opened the way for resolution of this dilemma. The early Earth, having higher stores of internal energy, had to dissipate it via extensive submarine hydrothermal systems and volcanism: even with today's diminished rates, this activity with no feedback would deplete the entire stores of atmospheric oxygen in about 10 million years by using oxygen for oxidation of Fe²⁺ to Fe³⁺ in basalts (and reduction of SO₄²⁻ to H₂S or by simple consumption of SO₄²⁻). The subsequent cooling of the Earth and the concomitant growth of continents meant that the main sink for oxygen was no longer the mantle and oceanic crust, but was instead the continents and their riverine systems⁸. Then as now, the bulk of oxidation/reduction coupling was confined to the carbon cycle, but the excess oxidative power on the early Earth was mopped up by the iron cycle, whereas now such a role is played by the sulphur cycle. If so, life mediated these processes but did not control the global balances. So it may have been the evolving planetary heat engine, reflected in the style of tectonic activity, not the biosphere, that led the ocean-atmosphere system over a time span of some 100 million years into a new steady state.

The contribution of Des Marais *et al.*⁴ provides additional support for this unconventional view of planetary evolution. Screening the organic carbon-isotope data from Proterozoic sediments for alteration, they note that the 'best' measurements define a unidirectional trend of decreasing $\Delta^{13}\text{C}$ with time. This may suggest about a twofold increase in the size of organic-carbon reservoir in the course of the Proterozoic and/or a progressive decline in atmospheric CO₂ levels⁹. In accord with the above 'upside-down' picture, the authors also note that intervals of increased burial of organic matter, causing release of oxidative power and drawdown of CO₂, appear to coincide with the episodes of enhanced tectonic activity.

The Earth system represents a nested hierarchy of cyclic populations, with smaller and faster cycles embedded in the larger ones. Evolution, then, may be seen¹⁰ as a response to an innovation whose impact is recorded on timescales

characteristic of given populations. A biological innovation of global significance should in any case saturate the entire environment in a geologically short time, even if thousands of cyclic generations were required for such a task. In the Precambrian geological record, with barely 10-million-year resolution, this must appear as an instantaneous stepwise jump in the signal. It follows that sloping trends extending over 100 million to a billion years, such as the carbon-isotope curve of Des Marais *et al.*, or the sulphur and strontium isotope records^{7,8}, cannot reflect biological innovations, but must be responses to tectonic cycles that operate on these timescales. If this is the case, the onset of a specific isotope signal does not necessarily herald the time of a biological innovation. For example, the onset of a large spread in the sulphide isotope record may result from coeval bacterial dissimilatory sulphate reduction, but does not necessarily mark its start. As long as the large-scale geological processes maintained the redox balance of the sulphur cycle, bacterial reduction — even if operative — was not required to restore the steady state.

It also follows that small and fast cycles cannot control overall balances much beyond their space or timescales. For example, should the rock cycle decide to store an additional 1 per cent of organic carbon every million years, the biosphere could at first respond by faster turnover of carbon in the extant oxidation/reduction couple, so maintaining a stable biomass. However, over 100 million years, the rock cycle will consume the entire organic carbon reservoir no matter what life does. The biosphere is indeed an immensely important player within the Earth system in terms of its catalytic, modulating and feedback functions, but only on biological, not geological, timescales. □

Ján Veizer is in the Geological Institute, Ruhr University, 4630 Bochum, Germany, and the Ottawa-Carleton Geoscience Centre, University of Ottawa, Ottawa, Ontario, Canada K1N 6N5.

The ACE of hearts

Theodore W. Kurtz

ON page 641 of this issue¹, Cambien *et al.* report data suggesting that a deletion polymorphism in the gene encoding angiotensin converting enzyme may be a risk factor for myocardial infarction (heart attack), particularly in those ordinarily thought to be at low risk for cardiovascular disease. Angiotensin converting enzyme is a key element of the renin-angiotensin system (see box) that has long been suspected to be involved in the pathogenesis of hypertension and cardiovascular disease. A number of drugs are already available that may influence the pathophysiological effects of this genetic variant on the cardiovascular system.

In a multicentre study designed to identify genetic determinants of myocardial infarction, Cambien and colleagues found that the frequency of a deletion

dial infarction. Earlier experimental studies had shown that ACE gene expression is increased in myocardial tissue after coronary artery occlusion⁴. Taken together, the papers by Cambien *et al.* and Pfeffer *et al.* have extraordinary clinical implications and will compel a host of follow-up studies. Will ACE genotyping be able to identify those patients in whom administration of an ACE inhibitor is most likely to prevent recurrent myocardial infarction? Will such genotyping be able to identify people in whom administration of an ACE inhibitor will be useful for the primary prevention of myocardial infarction? The potential clinical impact of these studies and many others that will flow from them cannot be overstated.

Second, unlike most association studies of coronary heart disease which

Angiotensinogen $\xrightarrow{\text{Renin}}$ Angiotensin I $\xrightarrow{\text{ACE}}$ Angiotensin II

Renin, a proteolytic enzyme, acts on angiotensinogen to generate the inactive prohormone angiotensin I. Angiotensin converting enzyme (ACE), a dipeptidyl carboxypeptidase, converts angiotensin I to the vasoconstrictor octapeptide, angiotensin II. ACE can also hydrolyse other physiological substrates and is responsible for the degradation of vasodilator kinins such as bradykinin. ACE inhibitors can decrease the formation of angiotensin II and the breakdown of bradykinin.

polymorphism in the gene for angiotensin converting enzyme (ACE) was significantly greater in patients with heart attack than in controls. This finding raises the possibility that genetic variation in or near the ACE locus may be a pathogenetic determinant. Although tobacco usage, high blood pressure, diabetes, obesity and lipid disorders (for example hypercholesterolemia) are well known risk factors for cardiovascular disease, investigators have always been puzzled by the fact that many people without such classic risk factors develop myocardial infarctions. Not only do Cambien *et al.* provide important molecular genetic clues to this puzzle, but their findings may also lead to improved strategies for the primary prevention of myocardial infarction.

Case-control analyses that reveal an association between a DNA polymorphism and a disease are quite common and are open to numerous criticisms². But the new study merits special attention for several reasons.

First, it comes on the heels of a dramatic clinical study in heart attack patients with asymptomatic left ventricular dysfunction, in which Pfeffer and colleagues³ found that administration of an ACE inhibitor not only decreased the risk for developing heart failure but also reduced the risk for recurrent myocar-

have focused on genes involved in lipid metabolism⁵, Cambien and colleagues targeted a candidate gene that can influence tissue and plasma levels of peptides affecting the heart independently of effects on lipid metabolism. Angiotensin II can promote growth of myocardial and vascular smooth muscle cells, neointimal hyperplasia after arterial wall injury, induce myocardial necrosis, stimulate activity in the sympathetic nervous system, and cause systemic and coronary vasoconstriction. Although angiotensin II can be generated independently of ACE⁶, *in vivo* administration of ACE inhibitors can reduce the synthesis of angiotensin II and attenuate these pathophysiological processes⁷. The beneficial effects of ACE inhibitors may also be related to their capacity to decrease ACE-induced degradation of kinins that have vasodilator and antiproliferative properties. In the new work, the association between the ACE deletion polymorphism and myocardial infarction seems to be independent of the usual lipid variables associated with increased risk for heart attack. It must be emphasized, however, that the mechanisms whereby ACE could be involved in the pathogenesis of myocardial infarction are complex and remain to be clearly defined.

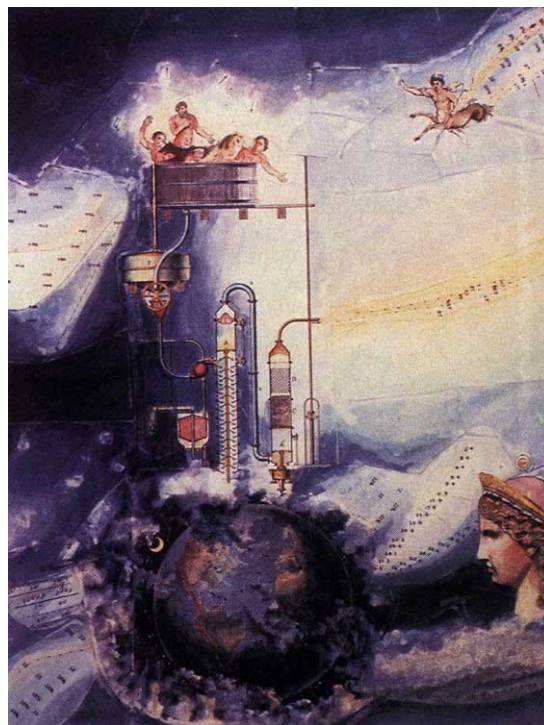
In genetic studies of spontaneous NATURE · VOL 359 · 15 OCTOBER 1992

- Schopf, J. W. (ed.) *Earth's Earliest Biosphere: Its Origin and Evolution* (Princeton University Press, 1983).
- Schopf, J. W. & Klein, C. (eds) *The Proterozoic Biosphere: Multidisciplinary Study* (Cambridge University Press, New York, 1992).
- Lovelock, J. *Gaia: A New Look at Life on Earth* (Oxford University Press, 1979).
- Des Marais, D. J., Strauss, H., Summons, R. E. & Hayes, J. M. *Nature* **359**, 605–609 (1992).
- Cloud, P. E. *Geol. Soc. South Africa, Annexe* **79**, 1–33 (1976).
- Broecker, W. S. *J. geophys. Res.* **75**, 3553–3557 (1970).
- Schidlowski, M., Hayes, J. M. & Kaplan, I. R. in *Earth's Earliest Biosphere: Its Origin and Evolution* (ed. Schopf, J. W.) 149–186 (Princeton University Press, 1983).
- Veizer, J. *Origin Life Evol. Biosphere* **18**, 13–39 (1988).
- Kasting, J. F. *Precamb. Res.* **34**, 205–229 (1987).
- Gregor, C. B. *Geochim. cosmochim. Acta* **56**, 2993–3000 (1992).

hypertension in animal models, ACE polymorphisms have been linked to the regulation of blood pressure^{8,9}. In the current study, and in sib-pair linkage studies of the ACE gene in human essential hypertension¹⁰, no relationship was found between ACE and increased blood pressure. So the increased risk for myocardial infarction in patients with the deletion polymorphism does not seem to be related to the level of systolic or diastolic blood pressure as typically measured in the brachial artery. However, given that the ACE gene can be expressed in the aorta, the relationship of ACE genotype to other haemodynamic determinants of cardiovascular disease such as aortic stiffness, aortic root pressure and pulse pressure may be worth looking at^{11,12}. The importance of the renin-angiotensin system in the genetic control of blood pressure has been underscored by recent linkage and association studies suggesting that molecular variants of the angiotensinogen gene contribute to the inheritance of essential hypertension¹³.

Third, the results of Cambien *et al.* derive further interest from the observation that, in hypertensive patients, increased levels of plasma renin activity are also associated with increased risk for heart attack, particularly in patients with a relatively low risk profile for cardiovascular disease¹⁴. It had previously been found that the plasma level of ACE in subjects homozygous for the deletion polymorphism was twice that in subjects without the deletion polymorphism¹⁵. Although measurements of circulating ACE or plasma renin activity are not reported by Cambien *et al.*, it is conceivable that a measurement profile that includes both plasma renin activity and ACE levels might be particularly useful for identifying individuals at increased risk for heart attack. Measurements of kinin levels or of angiotensin II, perhaps the ultimate

CHEMISTRY, especially its history, fuses with art and literature in the "Chemistry Imagined" exhibition currently running at the National Academy of Sciences, Washington DC. The picture reproduced here, *Greek Air*, is a characteristically surreal collage by Vivian Torrence in which drawings from old books are combined to illustrate particular propositions. The main theme though is that of collaboration — not only in collage as an art form and as a hallmark of much of contemporary scientific research, but because the exhibition is a joint endeavour between Torrence and Roald Hoffmann. Hoffmann, Nobel laureate in 1981 for his work on understanding the course of chemical reactions, provides prose and poetry to elaborate on and complement the pictures; among other titles of the collage-text pairs are *The Periodic Table*, *Electron Transfer*, *Air of Revolution* and *The Philosopher's Stone* (alchemy has its place). The exhibition runs in Washington until 6 January next year.



villain in this story, could also prove to be of interest.

However, it is difficult to obtain reliable measurements of plasma angiotensin II, kinins, ACE or renin activity in the clinical setting and, unlike genomic DNA sequence, these biochemical variables may change in response to cardiovascular disease and to changes in diet. Furthermore, measurements of circulating levels of these substances may not reflect their activities at the tissue level. Although still a matter of considerable controversy¹⁶, some investigators believe that the tissue renin-angiotensin and kallikrein-kinin systems may be of special functional significance. Given observations that the ACE insertion/deletion polymorphism may be a determinant of ACE levels in lymphocytes¹⁷, it is tempting to speculate that the ACE genotype also determines ACE activity in cardiovascular tissues. So the use of DNA-typing techniques to determine a risk profile for cardiovascular disease has both practical and theoretical appeal.

Finally, the current work is more than the usual association study in which a small group of disease patients is compared to a control group with totally unknown genetic background. Cambien *et al.* looked at fairly large sample sizes

and the relationship of ACE genotype to heart attack seems to hold across four different populations. Given possible questions about the strength of the statistical findings in the separate populations and the potential clinical implications of the results, however, it is essential that the findings should be independently replicated. Accordingly, the authors have been prudent in offering a concluding note of caution to their discussion.

Assuming that others do confirm that the ACE deletion polymorphism is associated with increased risk for heart attack, this will not prove that the ACE polymorphism itself is a culprit. Such positive association studies could be reflecting an effect of some other DNA sequence variant as much as 500 kilobase pairs away from the ACE insertion/deletion site¹⁸. For now, however, a molecular variant in the ACE gene seems to be the most likely suspect. In any case, Cambien *et al.* have dealt a very exciting hand to those interested in the molecular genetics of cardiovascular disease. □

Theodore W. Kurtz is in the Department of Laboratory Medicine, Box 0134, University of California San Francisco, San Francisco 94143-0134, USA.

- Cambien, F. *et al.* *Nature* **359**, 641–644 (1992).
- Cooper, D. N. & Clayton, J. F. *Hum. Genet.* **78**, 299–312 (1988).
- Pfeffer, M. A. *et al.* *New Engl. J. Med.* **327**, 669–677 (1992).
- Hirsch, A. T., Talsness, C. E., Schunkert, H., Paul, M. & Dzau, V. J. *Circ. Res.* **69**, 475–482 (1991).
- Berg, K. in *Genetic Approaches to Coronary Heart Disease and Hypertension* (eds Berg, K., Bulzhenkov, V., Christen, Y. & Corvol, P.) Vol. 1, 98–109 (Springer, Berlin, 1991).
- Urata, H., Healy, B., Stewart, R. W., Bumpus, F. M. & Husain, A. *Circ. Res.* **66**, 883–890 (1990).
- Swales, J. D. & Dzau, V. J. *Am. Heart J.* **123**, 1412–1413 (1992).
- Jacob, H. J. *et al.* *Cell* **67**, 213–224 (1991).
- Hilbert, P. *et al.* *Nature* **353**, 521–529 (1991).
- Jeunemaitre, X., Lifton, R. P., Hunt, S. C., Williams, R. R. & Lalouel, J. M. *Nature Genet.* **1**, 72–75 (1992).
- O'Rourke, M. *Hypertension* **15**, 339–347 (1990).
- Safar, M. E. J. *Hypertension* **7**, 769–776 (1989).
- Jeunemaitre, X. *et al.* *Cell* **71**, 169–180 (1992).
- Laragh, J. H. *Hypertension* **20**, 267–279 (1992).
- Rigat, B. *et al.* *J. clin. Invest.* **86**, 1343–1346 (1990).
- Swales, J. D. & Heagerty, A. M. *J. Hypertension* **5**, (suppl. 2), s1–s5 (1987).
- Costerousse, O., Allegrini, J., Lopez, M. & Alhenc-Gelas, F. *Biochem. J.* (in the press).
- Bodmer, W. F. *Cold Spring Harb. Symp. quant. Biol.* **51**, 1–13 (1986).

RÉSUMÉ

Beaming down

The dwarf star AU Microscopii (the name refers to its constellation, not its size) produces occasional flares of emission whose mechanism, which may also be relevant to solar flares, is little understood. Spectroscopic observations using the Hubble Telescope (B. E. Woodgate *et al.* *Astrophys. J.* **397**, L95–L98; 1992) of one revealed an increase in brightness on the redward side of the hydrogen Lyman- α emission line. This fits a proposed mechanism in which an energetic proton beam travels along a loop of magnetic field extending from the star. On the downward journey it interacts with neutral hydrogen, swapping electrons and generating Lyman- α emission with a small Doppler shift. Whether this was a freak or a characteristic event cannot yet be said.

Future attractions

DESIGNER magnets are the promise of H. Tamaki *et al.* (*J. Am. chem. Soc.* **114**, 6974–6979). The aim in trying to develop so-called molecular magnets is to combine magnetism with the material properties possible with modern synthetic chemistry. In general this work has concentrated on linear organic compounds, which are inherently one-dimensional. Tamaki *et al.* base their approach on a 'one-pot' synthesis using metal oxalate complexes to make three-dimensional examples. Two of the oxygens on each of the three oxalate ($\text{O}_2\text{C}_2\text{O}_2$) ions in a complex are bound to the central metal atom. The other three pairs bristle outwards so that, on addition of tetrabutylammonium to the dissolved metal oxalate, the complexes link up as a three-dimensional network to produce green microcrystals. Magnetic susceptibility measurements show that the microcrystals are ferromagnetic at low temperature.

Solar sterilization

CORRESPONDENTS to *The Lancet* continue to explore a low-technology way to help stem the rising tide of cholera in South America. *Vibrio cholerae* is largely transmitted through drinking water but can be inactivated by sunlight. Reporting on experiments in Ecuador, T. D. MacKenzie *et al.* (**340**, 367; 1992) suggest that the effectiveness of solar sterilization depends in part on altitude. T. Joyce *et al.* now weigh in with the results of experiments carried out in Dublin (**340**, 921; 1992). Their prescription is that the bottle used should be made of plastic not glass; and that water should be left to stand before being exposed to the Sun, to allow sediment to settle. Cloudiness in the water diminishes the effectiveness of the treatment and may, they say, have been responsible for the reduced disinfection seen at low altitude by MacKenzie and colleagues.

NEUROPSYCHOLOGY

Drawing upon the mind's eye

Jennifer M. Gurd and John C. Marshall

CAN an otherwise healthy adult plainly 'see' a familiar object (an umbrella, for example) but not recognize what it is? Surprisingly, yes. The condition, now known as associative visual agnosia, was first described by the Breslau neurologist Heinrich Lissauer in 1890¹ and a particularly clear-cut case is reported by Behrmann, Winocur and Moscovitch on page 636 of this issue².

These patients show a severe deficit of visual recognition despite perceptual abilities that seem to be sufficiently well-preserved to support identification by sight. They may thus be able to copy line drawings (or draw from a three-dimensional model) accurately enough to allow an observer to identify the drawing but are themselves unable to do so. By contrast, there is little or no difficulty in recognizing (and naming) the same objects by touch (or after tracing over visual forms with the finger)³. The lesion responsible is usually in the occipitotemporal regions of the left cerebral hemisphere.

In many such cases, drawing from memory is also seriously impaired⁴, even in patients who were skilled graphic artists prior to their brain lesion⁵. So it could be that the patients have lost from their pictorial memory the visual schemata of even quite common objects. This loss will rarely (if ever) be total. The recognition errors that patients make usually show some visual resemblance to the target object (C. K., the case of Behrmann *et al.*, named a pair of pliers as a clothes peg), but the distinguishing features for correct identification are unavailable. C. K. shows more convincingly than any previous case that the above account cannot apply to all instances of associative visual agnosia.

C. K. is grossly impaired in identifying written material, line drawings and real objects by sight, yet his copies (although constructed in a slavish, piecemeal fashion) are extremely accurate. By contrast, his drawing from memory is excellent: the examples shown in Fig. 1c (page 637) are detailed, well-proportioned and executed with an elegant, flowing line. As with his spontaneous handwriting in Fig. 1b, C. K. cannot identify his own drawings when they are shown to him some time later. His mental image of the objects he can 'see' (but not recognize) must accordingly be intact. This conclusion is borne out by C. K.'s flawless performance on tests of visual imagery; he can, for example, judge whether an imagined lower case letter contains an ascender or a descender and whether particular animals have

tails that are long or short in relation to their body size. As Behrmann *et al.* write, "No other case of such a clear dissociation between impaired object perception and intact imagery has been reported to date", although another recent paper, by Jankowiak and colleagues⁶, describes a patient who appears to be quite similar to C. K.

These cases contrast with patients who are unable to generate visual images 'in the mind's eye' but have reasonably intact recognition of what they see⁷. The double dissociation (intact imagery with impaired recognition versus impaired imagery with intact recognition) suggests that functionally (and perhaps anatomically) distinct routes to a central store of visual schemata can be selectively damaged.

What, then, might account for identification failure despite apparent 'seeing'? According to one hypothesis, higher levels of perception are not fully intact in the associative visual agnosias. Observe C. K.'s copy of an array of a circle and two diamonds (Fig. 1a on page 637). It is obvious from the stroke pattern of C. K.'s copy that he has not seen these forms as a circle and two diamonds. There is a failure of pictorial parsing despite the fact that the final outcome is an accurate copy. If access to pictorial memory is based on a correctly parsed visual input, C. K.'s recognition failure would follow from his failure to 'see' the relationships between the component elements of the stimulus pattern⁸.

An alternative (or complementary) hypothesis, proposed by Jankowiak *et al.*⁶, suggests that deficient lateral inhibition between the neural representations of visually similar objects could be responsible for the impairment. If perceptual input triggered a small set of rival interpretations, a failure of the lateral inhibitory feedback that normally amplifies small differences between visual configurations could result in misidentifications that preserved the overall shape of

1. Lissauer, H. *Arch. Psychiat.* **21**, 222–270 (1890).

2. Behrmann, M., Winocur, G. & Moscovitch, M. *Nature* **359**, 636–637 (1992).

3. Landis, T., Graves, R., Benson, F. & Hebben, N. *Psychol. Medicine* **12**, 515–531 (1982).

4. Mehta, Z., Newcombe, F. & De Haan, E. *Neuropsychologia* **30**, 645–655 (1992).

5. Wapner, W., Judd, T. & Gardner, H. *Cortex* **14**, 343–364 (1978).

6. Jankowiak, J., Kinsbourne, M., Shalev, R. S. & Bachman, D. L. *Cog. Neurosci.* **4**, 119–131 (1992).

7. Basso, A., Bisiach, E. & Luzzati, C. *Neuropsychologia* **18**, 435–442 (1980).

8. Kanizsa, G. *Organization in Vision: Essays on Gestalt Perception* (Praeger, New York, 1979).

9. McCarthy, R. A. & Warrington, E. K. *Cognitive Neuropsychology: A Clinical Introduction* (Academic, New York, 1990).

10. Young, A. W. *Phil. Trans. R. Soc.* **B335**, 47–54 (1992).

the target stimulus.

This latter interpretation might be thought incompatible with the most remarkable of C. K.'s preserved perceptual abilities. He has no difficulty in recognizing famous faces or in matching different views of unfamiliar faces. This preservation of face recognition in a small minority of patients with severe object agnosia has previously been noted⁹, but its significance has perhaps not been fully appreciated. The central controversy in the study of face recognition concerns whether or not there is a specific brain module dedicated to this process¹⁰. Many neuropsychologists have argued that faces are merely the most extreme example of stimuli where small physical differences make a large psychological difference. They therefore interpret cases of prosopagnosia (im-

paired familiar face recognition) without object agnosia as reflecting no more than the difficulty of perceiving the very slight differences between faces that are crucial to the identity of the individual. By contrast, the recognition of a chair as a chair does not depend upon precise discrimination of the relative sizes and positions of component parts.

C. K., however, cannot see the fine distinctions that differentiate between a dart and a feather duster, whereas he can distinguish between very similar faces. This pattern of performance constitutes strong evidence for a dedicated face recognition module. □

Jennifer M. Gurd and John C. Marshall are in the Neuropsychology Unit, Department of Clinical Neurology, Radcliffe Infirmary, Oxford OX2 6HE, UK.

gel method was shown to be especially useful for making such materials as alumina/silicon carbide and alumina/metal nanocomposites. This kind of approach is particularly popular in the United States, maybe as a result of pressure from funding agencies which have themselves been influenced by a call from the US National Academies of Science and Engineering for more research emphasis on materials synthesis².

These novel methods can produce uniform nanomaterials or composites of various kinds. Among the composites are 'engulfed' nanodispersoids made by biomimetic colloid methods (I. A. Aksay, Princeton University) and unusual nanoparticles of metals inserted in zeolite cages, as described by one of the master researchers in the field (R. Siegel, Argonne National Laboratory, Illinois).

Another category of experiment which is rising in prominence is the study of free (unconsolidated) nanostructured powders. At the meeting there was especial interest in the change of melting temperature and lattice parameter as a function of particle size in both metallic and compound powders; the sharp depression of melting point was repeatedly confirmed. A curious aspect was that, in some studies, repeated melting and re-solidification revealed no change in the shift of melting temperature, whereas in others (aluminium, for example) there was a changed melting temperature after one cycle, apparently because the lattice strain in the original nanopowder had affected the melting process.

MATERIALS SCIENCE

Nanostructures come of age

Robert W. Cahn

Two years ago I commemorated in these pages¹ an invitation symposium on nanostructured materials (that is, materials consisting of particles measuring a few nanometres in one, two or all dimensions), an event which revealed the highly variegated nature and potential uses of such materials. Last month the first open and international meeting* on the topic took place at Cancún in remote Yucatán, Mexico, and both the geographical origin of the participants and the scientific topics were even more variegated. Research interest in this intriguing subject has spread rapidly, and the meeting reflected in particular the pronounced advances seen in methods of synthesis. It is no accident that those advances are being accompanied by the emergence of a range of highly unexpected practical applications.

In 1990, almost all the researches reported were based on 'cluster-assembled' materials, metallic or ceramic; that is to say continuous solids made by consolidating nanostructured powders, themselves made by evaporation followed by condensation in gas at relatively high pressure. This method, the original one invented by the founder of the field, Herbert Gleiter, produces only sub-gram quantities; consequently, many experimental measurements, of mechanical properties in particular, are delicate and difficult. At the previous meeting, some experimenters (B. H. Kear, for example) reported alternative

methods, most notably chemical reactions under conditions which favour the formation of unusually fine precipitates. This time, Kear and McCandlish (Rutgers University, New Jersey) described a novel 'spray-conversion process' for making nanocomposites such as the WC/Co 'hard metal' from organometallic precursors; the process allows these composites to be produced in quantity, and in the nanoform they make

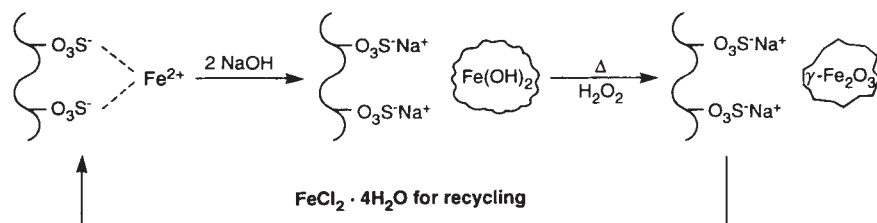


FIG. 1 Diagram of the $\gamma\text{-Fe}_2\text{O}_3$ precipitation process using Fe(II) in an ion-exchange resin. (From ref. 4.)

more wear-resistant cutting tools than in the normal coarse form.

A number of other modes of chemical (or non-chemical) synthesis were reported to be useful for making nanostructures, including a range of nanocomposites: gas-phase combustion synthesis, crystallization of glasses, sol-gel processing and other colloid-based procedures, mechanical milling and electrodeposition, and the use of ion-exchange resins, were all described at Cancún. It is clear that the original approach is well on the way to being overtaken by these newer methods, which can produce much larger quantities of both free powders and continuous materials. The sol-

No doubt the mode of preparation makes a difference in this connection. An especially stimulating talk came from a physicist (L. J. de Jongh, University of Leiden, Netherlands) who spoke about metallic clusters stabilized by complex chemical ligands; such clusters have large 'magic numbers', that is cluster sizes which are particularly stable and therefore form preferentially. This kind of chemically stabilized cluster is such an important topic that it has spawned its own journal, the *Journal of Cluster Science* (recently reviewed in *Nature's* new journals issue³). Here again there are undoubted practical implications.

M. L. Trudeau *et al.* (Hydro-Québec

* First International Conference on Nanostructured Materials, Cancún, Mexico, 21–26 September 1992. Proceedings to be published as a special volume of *Nanostructured Materials*.

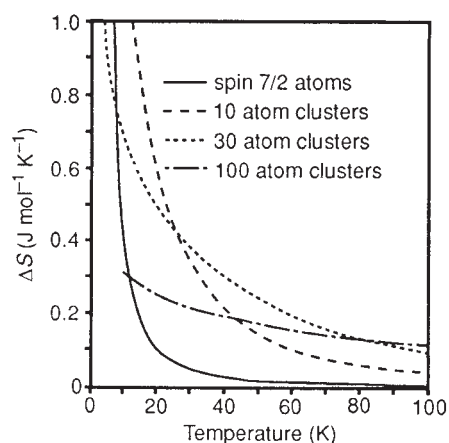


FIG. 2 Entropy change of a system of spin- $7/2$ Gd atoms induced by removal of a 1 tesla field for individual spins and clusters of 10, 30 and 100 spins. The 10-atom clusters have the optimum size at 15.3 K, the 30-atom clusters at 46 K. (From ref. 5.)

and McGill University) showed that FeTi, a standard intermetallic used for storing hydrogen as hydride, can both store more hydrogen per unit mass and is more easily conditioned (that is, needs fewer hydrogenation cycles to attain its optimum performance) than coarse-grained material; Trudeau and colleagues opine that the oxide layer is of a different nature on nanostructured powder.

A hardy perennial in the study of consolidated nanostructured materials is the curiously strong resistance of the very small grains, typically 10–20 nm across, to grain growth during the sintering process used to consolidate the powder. Several papers at Cancún claimed to show that the key factor was the size, shape and separation of the residual pores. As long as porosity is connected, grain growth is inhibited; once pores become separated, grain growth becomes easier unless the pores remain very small and numerous. But J. Weissmüller (University of Saarbrücken, Germany) puts a quite different perspective on this issue. He starts from his own experimental measurements (by analytical electron microscopy with nanometric beam diameter), showing that grain boundaries, for instance in Ti/Pd nano-alloys, are enriched in one constituent. He then demonstrates, by rigorous thermodynamic argument, that such enrichment can entirely remove any driving force for grain growth by leaving the enriched grain boundary with an effectively negative specific boundary energy. This analysis (echoed in another paper by H. J. Fecht, Augsburg University) will certainly attract wide attention.

Finally, several striking papers were devoted to magnetic properties. R. F. Ziolo (Xerox Corporation, Webster, New York), jointly with academic input by several other scientists, has found a

way of making a transparent ferro-magnetic nanocomposite, containing $\gamma\text{-Fe}_2\text{O}_3$, which is much more strongly magnetic than the only earlier transparent ferromagnetic materials, FeBO_3 and FeF_3 . A preliminary account of this work was published in July⁴. The key to the process is the use of a standard sulphonated ion-exchange resin, treated as shown in Fig. 1. The optical transparency is good, especially above 600 nm wavelength; this characteristic is clearly linked to the small size of the particles relative to the light wavelength. The embrittled resin beads can be broken up and the magnetic nanoparticles dispersed to make a ferrofluid (needing only short milling times), and the cast and dried ferrofluid forms soft, flexible films that adhere well to many substrates. Uses in magneto-optical systems and light modulators are suggested.

The other strikingly original magnetic application of a nanostructured material was presented by R. D. Shull *et al.* (National Institute of Standards and Technology, Maryland). They showed, both theoretically and experimentally, that ferromagnetic or preferably superparamagnetic clusters, typically containing 10–30 atoms each, in a non-magnetic matrix, are far more effective for magnetic cooling (by means of an adiabatic magnetization/demagnetization cycle) than are conventional magnetic materials. In the past magnetic cooling has been restricted to temperatures below about 20 K. But by careful choice of material parameters and cluster size, highly effective magnetic cooling can now be obtained at temperatures as high as 100 K. Figure 2 shows the entropy change on demagnetization (closely related to the attainable ΔT per cycle) as a function of cluster size in Gd. The underlying theory is too complex to outline here, but is clearly set out in the paper from which the figure is taken⁵ and other papers by the group.

There were many other worthwhile topics covered at the meeting. Not least of them were metallic multilayers (nanostructures small in one dimension only) and fullerene nanotubes⁶ (nanostructures small in two dimensions, as described by S. Iijima, NEC, Japan). But the main message to emerge was a broad and simple one — a major new category of materials has firmly arrived. □

Robert W. Cahn is in the Department of Materials Science and Metallurgy, University of Cambridge, Pembroke Street, Cambridge CB2 3QZ, UK.

1. Cahn, R. W. *Nature* **348**, 389–390 (1990).
2. Cahn, R. W. *Adv. Mater.* **2**, 206–207 (1990).
3. Cahn, R. W. *Nature* **359**, 456–457 (1992).
4. Ziolo, R. F. *et al. Science* **257**, 219–223 (1992).
5. McMichael, R. D. *et al. J. Magn. Magn. Mater.* **111**, 29–33 (1992).
6. Iijima, S. *Nature* **354**, 56–58 (1991).

DAEDALUS

High vacuum

LAST week Daedalus presented his 'actively rigid' hollow tube. It carries an internal laser or electron beam. If the tube is bent out of the straight, the beam impinges on the inner face of the bend. The material at that point heats up, expands, and straightens the tube again. Thus protected against buckling, it can withstand enormous compression or deformation.

DREADCO engineers are now trying this idea in two dimensions. They are laminating flat sheets of transparent material between two opaque faces, and launching a laser beam into the transparent central region. If the sheet deflects in any way, the beam hits the inner surface of the deflection and straightens it again. The beam has to be diverged into a sheet form itself, or scanned rapidly enough through the inner volume to correct any deflection almost instantly. The result is a thin but actively rigid panel, proof against all bending and buckling.

Such light, flat panels could replace heavy walls, roofs and partitions in many types of building. But Daedalus would like to make them in curved sections as well. He plans to give the transparent central sheet a refractive index gradient from one face to the other. The laser beam will then travel in a curve, and the panel will follow it. Actively rigid domes, arches, tanks and pipes of the new material could be as strong as thick steel, yet almost as thin and light as paper. A complete actively rigid hollow sphere would be so strong that it might even be evacuated without collapsing.

Vacuum, notes Daedalus, is lighter than hydrogen, and much cheaper. Accordingly, an evacuated lightweight actively rigid sphere would form a novel vacuum balloon. Unlike all other balloons, it would be independent of gas supplies. It could vary its lift at will, simply by letting air into its envelope or pumping it out again.

Bolder still, Daedalus is dreaming of a cylindrical vacuum balloon in the form of a tall, actively rigid evacuated chimney. Unlike conventional chimneys and masts, it would be buoyant and self supporting in the atmosphere. Winds could not deform or topple it: at worst they might push it harmlessly out of the vertical. It could rise many kilometres through the atmosphere, and could even open at the top into the near vacuum of aerospace. A telescope looking up the bore could then see the stars free of atmospheric turbulence. Equipment at the top could conduct space experiments, and relay telephone and TV signals over a huge area. A feeble downflow of thin high-altitude air could be pumped steadily out at the bottom and sampled for ozone and halocarbons.

David Jones

Some like it hot . . .

SIR — Wehner, Marsh and Wehner attribute the lower limit of the foraging temperature range of the Saharan silver ant, *Cataglyphis bombycina*, to be determined by predation pressure exerted by the lizard, *Acanthodactylus dumerili*¹. This assumption may not be justified, as the authors do not show that the ant-eating lizard does in fact exert sufficient selective pressure on the ant population, and especially considering that the ants, when experimentally released at temperatures below their natural foraging window, fell victim to the lizards within "less than five minutes". The relative distribution of the two species as a function of the daily temperature could be independent phenomena and not causally related. It is also not clear how "during the hottest time of a summer day, this lizard concentrates on *C. bombycina* as its potential prey . . .", in view of the fact that "the lizard . . . retreats underground at about the temperature at which the ants' foraging activity begins".

We believe that the lower limit of this remarkable narrow foraging window may instead be shaped by interspecific competition for food resources. Differences in foraging patterns among sympatric species of ants can serve in temporal partitioning of significant resources². An example is the black ant, *Monomorium minimum*, in the grasslands of Massachusetts, which is able to forage at higher temperatures than its closest competitors, *Lasius neoniger*, *Myrmica americana* and *Tetramorium caespitum*³. Evidence for competition has been found for such disparate desert dwellers as birds, rodents, lizards and ants⁴, an indication that interspecific competition may well be widespread in such arid environments. *C. bombycina* may thus turn out to be just an efficient forager, scouring the hot desert floor for dead and dying prey at midday temperatures that no other potential arthropod competitor has yet been able to exploit.

The physical constraints imposed on the silver ant by the need of thermoregulation is also not very unusual. Ground-living desert arthropods and reptiles maintain their preferred body temperature (eccritic temperature) by regularly alternating between sites where they gain heat and those where they shed excess heat⁵. In many of these animals, the eccritic temperature is close to the upper lethal temperature. It has been postulated that the evolution of such high eccritic 'set points' may help individuals to perform strenuous heat-generating activity for extended periods of time, especially if heat dissipation is slow under conditions of high ambient temperatures⁶. The characteristic eccritic

temperature of *C. bombycina* may thus allow this species to withstand prolonged periods of heat, so enhancing its foraging efficiency; this idea is reinforced by the authors' observation that individuals can even afford to rest for as long as 75% of their search time.

Many ant species, including several desert taxa², can shift their time of foraging as an adjustment to the vagaries of the environment. For example, the red harvester ant, *Pogonomyrmex barbatus*, from the Sonoran desert not only shows age-related changes in colony organization and activity patterns, but also a more important, complex moment-to-moment regulation of task-performance in response to short-term environmental changes⁷. Is the foraging pattern of *C. bombycina* sensitive to

seasonal, or even short-term, variations in food availability? Although Wehner *et al.* have put forward an interesting interpretation, a more extensive horizontal and longitudinal study may be necessary to dissect out the true environmental correlates of the unusual foraging behaviour of the silver ant.

Anindya Sinha

K. Chandrashekara

Centre for Ecological Sciences,
Indian Institute of Science,
Bangalore 560 012, India

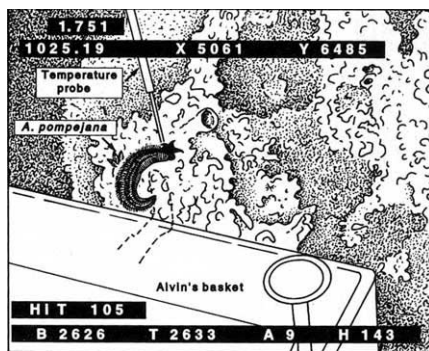
1. Wehner, R., Marsh, A. C. & Wehner, S. *Nature* **357**, 586–587 (1992).
2. Hölldobler, B. & Wilson, E. O. *The Ants* (Springer, Berlin, 1990).
3. Adams, E. & Traniello, J. *Oecologia* **51**, 265 (1981).
4. Wagner, F. H. & Galetz, R. D. in *Arid-land Ecosystems: Structure, Functioning and Management Vol 2* (eds Goodall, D. W., Perry, R. A. & Howes, K. M. W.) 51–83 (Cambridge University Press, 1981).
5. Cloudsley-Thompson, J. L. *Ecophysiology of Desert Arthropods and Reptiles* (Springer, Berlin, 1991).
6. Heinrich, B. *Am. Nat.* **111**, 623–640 (1977).
7. Gordon, D. M. *Anim. Behav.* **35**, 833–843 (1987).

. . . and some even hotter

SIR — Wehner *et al.*¹ recently reported observations of a desert ant (*Cataglyphis bombycina*) that experiences exposure to high-temperature while foraging under the midday sun, making this ant the most 'thermophilic' terrestrial animal, with a body temperature sometimes exceeding 50 °C. Apart from the fact that *Eremogarypus perfectus*, a pseudoscorpion from the Namib desert, still holds the terrestrial high-temperature record of 65 °C (ref. 2), the new result led us to ask whether the documented thermo-

tolerance of aquatic eukaryotes exceeded that of terrestrial representatives. The highest temperature at which growth of several aquatic taxa can occur is thought to be 50 °C for invertebrates, 56 °C for protozoa, 60 °C for algae and 62 °C for fungi (the highest value recorded for a eukaryote)³. It has been proposed that animals are limited to less than 50 °C, and that 60 °C is the thermal barrier beyond which eukaryotes cannot complete their life cycle because of their inability to form functional thermostable organellar membranes³. But the eggs or cryptobiotic forms of certain invertebrates can survive exposure to temperatures approaching boiling at atmospheric pressure^{4,5}. Our recent observations on deep-sea hydrothermal vent fauna show that certain marine organisms have thermotolerances that far surpass those of their terrestrial counterparts.

Alvinellid polychaete worms dwell in whitish parchment-like tubes secreted in the hottest part of the hydrothermal vent ecosystems, on the wall of chimney-like mineral structures called smokers. Here,



Main picture, photograph of a video taken from DSRV *Alvin* in hydrothermal vents of the deep Eastern Pacific, showing an *A. pompejana* standing on a substrate measured at 105 °C. Top, sketch to clarify the position of the worm and the probe.

the gradient between the sea water (2 °C) and the hot (350 °C) toxic hydrothermal fluid is so extreme that spatial measurements need to be accurate to the centimetre. Unlike the other prominent members of these communities, such as vestimentiferan tube worms and brachyuran crabs, experimental study of the alvinellids is impossible as they rarely reach the surface alive. One of the most challenging issues is that of the temperature range that the two *Alvinella* species are able to withstand. Previous preliminary temperature measurements⁶ revealed the temperature range within a colony to be dramatic: 20 °C at the tube openings, 100 °C 10 cm into the colony, and more than 250 °C 10 cm deeper. Although it was unclear exactly where the worms reside in the colony structure, later observations have reported individuals living around 50 °C (ref. 7).

Our most recent results, obtained in October 1991 and April 1992 during the joint French-US HERO (Hydrothermal Environment Research Observatory) expeditions at 13° N on the East Pacific Rise, are based on discrete temperature measurements made with the probes of the submersibles *Nautile* and *Alvin*, and were simultaneously recorded on video film. Nine measurements made on the animals' branchial funnels (at the opening of the tubes) were in the range of 20–45 °C. Six measurements were then made 2–3 cm inside tubes, four of these are 40–50 °C and the other two 70–80 °C. The most astonishing observation was a dislodged live *Alvinella pompejana* that, for a few minutes, coiled around the tip of *Alvin*'s high-temperature probe, which simultaneously recorded 105 °C. The worm then returned to the chimney surface, displaying a behaviour similar to that of other worms when out of their tubes. It has been previously reported that *A. pompejana* could withstand high temperatures, but this is the first time that such an observation has been recorded on video (see figure).

This extreme temperature clearly exceeds the physiological optimum suggested for *Alvinella* species (20–50 °C) by biochemical criteria (haemoglobin O₂ affinity⁸, malate dehydrogenase kinetics⁹, collagen melting temperatures¹⁰ and mitochondrial respiration¹¹).

However, our observations raise the question of the upper temperature limit of these organisms, perhaps the highest known for invertebrates and possibly eukaryotes.

Pierre Chevaldonné

Daniel Desbruyères

IFREMER Centre de Brest,
B.P. 70, 29280 Plouzané, France

James J. Childress

Marine Science Institute,
University of California,
Santa Barbara, California 93106, USA

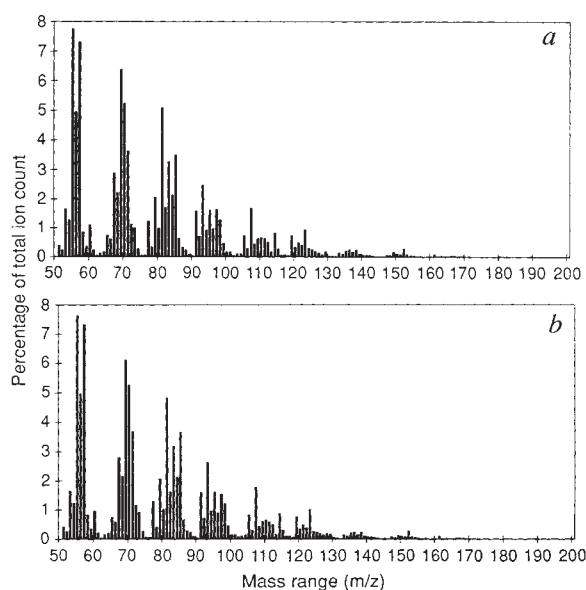
Neural networks and olive oil

SIR — Virgin olive oil is the oil extracted by purely mechanical means from sound, ripe fruits of the olive tree (*Olea europea* L.). Such oils with a free fatty acid

(PyMS)^{2,3} with multivariable data analysis using artificial neural networks⁴ has permitted us to effect a rapid assessment of the adulteration of extra virgin olive oils with various seeds oils.

Two sets of samples were prepared in G.B.'s laboratory, each consisting of 12 samples of various extra-virgin olive oils plus 12 samples variously adulterated with 5–50% of soya, sunflower, peanut, corn or rectified olive oils. The experiment was performed double-blind, such that the identities of the second set were not known to any of us. PyMS was performed at 530 °C using a Horizon Instruments PyMS-200X machine³; two typical spectra are shown in the figure, where it is clear that their distinction by eye is difficult.

We trained an artificial neural network consisting of an input layer of the 150 normalized ion intensities with mass: charge in the range 51–200 and one hidden layer of eight nodes, using the standard back-propagation algorithm⁴ as implemented in the NeuralDesk package (Neural Computer Sciences, Totton, Southampton), coding virgin oils as 1, non-virgins as 0. When the network had trained (r.m.s. error < 0.001), we tested it on the unknowns. When the code was broken, it transpired that the network had correctly assessed each oil. In a typical run, the virgins were assessed with a code of 0.99976 ± 0.000146 (range 0.99954 – 1.00016) and the non-virgins with a code of 0.001079 ± 0.002838 (range 0.00026 – 0.01009). We conclude that the combination of Curie-point



Pyrolysis mass spectra of a virgin olive oil (a) adulterated with 5% soya oil (b).

content (in terms of oleic acid) below 1% are termed 'extra virgin', whereas those with good flavour but greater acidity may be graded as 'fine' or 'semi-fine'. Lower grades, including those that have been subjected to refining, are called 'lampante' or 'pure'. Olive oil is considered to contribute significantly to the nutritional and health benefits of Mediterranean-type diets and, uniquely among vegetable oils, the flavour of olive oil is best enjoyed without refining. Olive oil therefore commands a higher price than other vegetable oils, and these and other properties mean that there is a great temptation to adulterate olive oils with other seed oils¹. Although various methods have been proposed for the detection of olive oil adulteration¹, none has found widespread usage. We wish to report here that a combination of Curie-point pyrolysis mass spectrometry

networks constitutes a powerful approach to the assessment of food adulteration.

Royston Goodacre

Douglas B. Kell*

Department of Biological Sciences,
University of Wales,

Aberystwyth, Dyfed SY23 3DA, UK

Giorgio Bianchi

Istituto Sperimentale per la Elaiotecnica,
Contrada "Fonte Umato" n37,
65013 Città S. Angelo,
Pescara, Italy

* To whom correspondence should be addressed.

1. Wehner, R. et al. *Nature* **357**, 586–587 (1992).
2. Heurtaut, J. & Vannier, G. *Acta Zool. Fennica* **190**, 165–171 (1990).
3. Brock, T. D. *Thermophilic Microorganisms and Life at High Temperatures* (Springer, New York, 1978).
4. Hinton, H. E. *Nature* **188**, 336–337 (1960).
5. Carlisle, D. B. *Science* **161**, 279–280 (1968).
6. Desbruyères, D. et al. *C. r. hebdo. Séanc. Acad. Sci., Paris* **295**, 489–494 (1982).
7. Baross, J. A. & Deming, J. W. *Bull. Biol. Soc. Wash.* **6**, 355–371 (1985).
8. Toulmond, A. et al. *Biol. Bull.* **179**, 366–373 (1990).
9. Dahlhoff, E. & Somero, G. N. *J. exp. Biol.* **159**, 473–487 (1991).
10. Gailli, F. et al. *J. molec. Biol.* **221**, 209–223 (1991).
11. Dahlhoff, E. et al. *Physiol. Zool.* **64**, 1490–1508 (1991).

1. Kiritsakis, A. & Marakakis, P. in *Essential Oils and Waxes* eds. Linskens, H. F. & Jackson, J. F. 1–20 (Springer, Heidelberg, 1991).
2. Meuzelaar, H. L. C., Haverkamp, J. & Hileman, F. D. *Pyrolysis mass spectrometry of Recent and Fossil Biomaterials* (Elsevier, Amsterdam, 1982).
3. Aries, R. E., Gutteridge, C. S. & Ottley, T. W. *J. analyt. Appl. Pyrol.* **9**, 81–98 (1986).
4. Rumelhart, D. E. et al. *Parallel Distributed Processing* (MIT, Cambridge, 1986).

Newtonian rigour

William R. Shea

The Investigation of Difficult Things: Essays on Newton and the History of the Exact Sciences. Edited by Peter M. Harman and Alan E. Shapiro. Cambridge University Press: 1992. Pp. 531. £90, \$175.

THE title of this collection of essays in honour of T. D. Whiteside is taken from the last Query of the *Opticks* where Newton refers to mathematics and natural philosophy as the study of "the investigation of difficult things". Like Newton, Whiteside has never shirked the investigation of difficult things, as the most cursory inspection of his impressive list of publications at the beginning of this book confirms. The 20 historians of science who contributed to this Festschrift can also be relied upon not to shun technical problems; indeed, they may be said positively to enjoy grappling with the finer points of historical scholarship from lunar velocity in the Ptolemaic tradition (Bernard R. Goldstein) to Poincaré's topological dynamics (Jeremy Gray), not to mention shadow measurement by Kepler (N. M. Swerdlow) or the reason why Gabriel Stokes never wrote on optics (Jed Z. Buchwald).

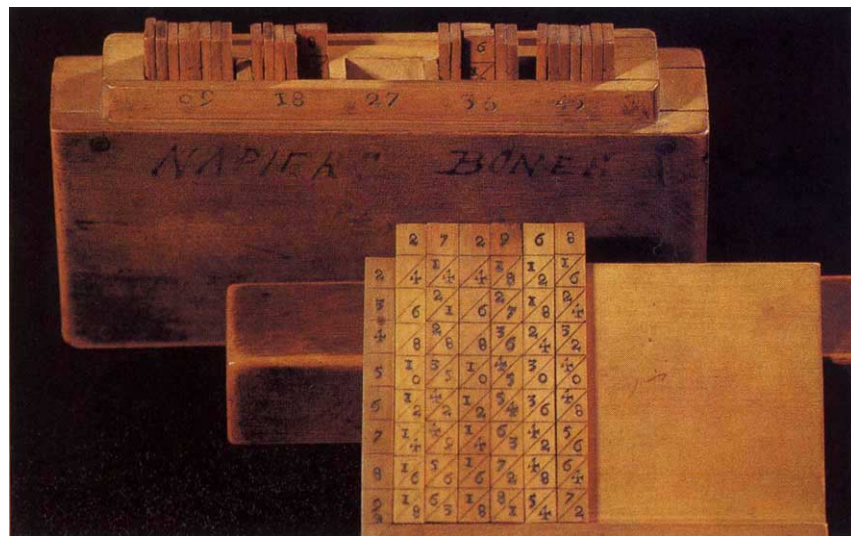
Whiteside's outstanding contribution to the history of science, and one of the landmarks of twentieth-century scholarship, is his critical edition of *The Mathematical Papers of Isaac Newton*, published in eight volumes between 1967 and 1981. When Newton died intestate in March 1727, his papers passed into the hands of his niece Catherine Barton who had married John Conduitt, Newton's successor at the Royal Mint. In 1740, their only child married John Wallop, whose father later became the first Earl of Portsmouth. In 1872, the fifth Earl entrusted all of Newton's papers to the University of Cambridge where they were catalogued, without undue haste or rigour, by a committee that included Stokes and John Couch Adams for the section on mathematics. On completion of this task in 1888, the portion of the papers regarded as scientific was presented by the Earl of Portsmouth to the university and deposited in the university's library. The rest of the collection was returned to Lord Portsmouth, his heirs offering it for sale at Sotheby's in 1936. The more than three million words consisted mainly of Newton's alchemical, theological, historical and Mint papers, as well as some important correspondence. It is interesting to note that the alchemical papers were competently classified in the 1888 catalogue, but that their content struck no sympathetic chord in G. D. Liveing, the professor of

chemistry who had been a member of the committee and is reputed to be the author of the verdict: "Newton's manuscripts on Alchemy are of very little interest in themselves." Thanks to John Maynard Keynes, many of the alchemical manuscripts were to return to Cambridge. The theological material eventually found its way to Jerusalem. P. E. Spargo writes entertainingly in this book about the 1936 sale of Newton's manuscripts and the horse-trading that went on between Keynes and the oriental scholar and book collector Abraham Shalom Ezekiel Yahuda. One of the alchemical manuscripts that did not return to Cambridge but strayed as far as Stanford, *De Scriptoribus Chemicis*, is the object of a detailed study by Karin Figala, John Harrison and Ulrich Petzold, who use it to reconstruct the contents of Newton's alchemical library, which contained well over 125 titles.

The Portsmouth Collection was still languishing largely unread when Whiteside arrived in Cambridge in 1956, with a bachelor's degree in French and Latin from the University of Bristol. During his research for his doctoral dissertation, "Patterns of Mathematical Thought in the Later Seventeenth Century" (which he wrote in 29 days in the spring of 1959), Whiteside was enthral-

led by the "depth and luxuriance" of Newton's unexplored mathematical papers. He was less enthusiastic about their extremely disordered state. Not only did they lack chronological sequence, but frequently the sheets of a single work were separated and scattered throughout the collection. Whiteside's first labour was to date this vast material. He tells us in the introduction to his edition of *The Mathematical Papers*, "we came ultimately to trust our ability to fix the date of composition of an unexamined portion of autograph manuscript accurately by sight to within half a dozen years (and sometimes even more narrowly still)". Alan E. Shapiro, who has been working on the chronological ordering of Newton's optical papers, is less sanguine about the possibility of determining the date of documents by the handwriting. In this book he contributes an interesting paper on the use of watermarks in dating documents, concluding that they cannot provide the magic key that unlocks the date of Newton's papers to two or three years, as he had initially hoped.

Whiteside's more creative task was to find a way of presenting Newton's research so that it would enlighten the scholarly community. His remarkable knowledge of the mathematics of the period has enabled him to bring Newton's mathematical thinking back to life and to place it solidly within the context of seventeenth-century mathematics. Whiteside's notes and extended commentaries are treasure-houses that contain a wealth of information about a surprising variety of aspects of science in Newton's day. Whiteside has also been unstinting in his help to scholars who have called on him for advice, a fact



Napier's bones — John Napier (1550–1617), renowned for his invention of logarithms, was also responsible for this aid to calculation, a multiplication table sliced up into moveable columns. The picture is taken from *Making Of the Modern World*, which illustrates 100 key discoveries using collections from the Science Museum, London. Published on 5 November by Murray, £17.95.

mentioned by several contributors in this volume, including, for instance, J. Bruce Brackenridge when addressing a moot point in his essay on the role of curvature in the genesis of Newton's dynamics. The influence of Whiteside on today's editorial practice can be seen in Harman's suggestive essay on James Clerk Maxwell and Saturn's rings which, he tells us, is a by-product of his continuing edition of Maxwell's scientific letters and papers.

Less obviously influenced by Whiteside but akin in scholarly spirit are Curtis Wilson's lucid discussion of Euler's views on action at a distance; Dominico Bertoloni Meli's perceptive comments on the problem of free-fall around 1800; Ronald Gowing's analysis of the work of Cotes and Varignon on spirals, and Lenore Feigenbaum's refreshingly nontechnical study of the European mathematical community in the first two decades of the eighteenth century. French mathematics in the seventeenth century is well represented with essays by Henk J. Bos on Descartes' solution of Pappus's problem, and Emil A. Fellmann (writing in German) on the Jesuit Honoré Fabry. Several contributions deepen our understanding of Newton's own achievements: Alan Gabby shows how Newton transformed the traditional notion of mechanics; D. H. Fowler provides an erudite footnote to Newton's theory of the resistance of fluids; and Zev Beckler ponders the ontology behind Newton's concepts of force and inertia.

Two contributions by senior scholars who encouraged Whiteside in his career deserve special mention. These are A. Rupert Hall's judicious reappraisal of the sources of Newton's conception of absolute space and absolute time, and I. Bernard Cohen's analysis of the reaction of the scientific community to Newton's ideas in the light of the reviews of the *Principia* that appeared in the main scientific journals of the period, namely *Acta Eruditorum*, *Bibliothèque Universelle*, *Journal des Sçavans* and *Philosophical Transactions*. The most complete account was given in *Acta Eruditorum* and prompted Leibniz to publish his own very different explanation of the cause of planetary motion.

The *Investigation of Difficult Things* is a fitting tribute to a distinguished editor and a remarkable scholar. It will find its place in any good library for the history of science next to the eight volumes of *The Mathematical Papers of Isaac Newton*, which have become an essential reference for anyone interested in the history of the scientific revolution. □

William R. Shea is in the Centre for Medicine, Ethics and Law, McGill University, 2020 University, Suite 2410, Montreal, Quebec, Canada H3A 2A5.

Large cast, but no plot

Gregory A. Petsko

Macromolecular Structures 1991/1992.

Edited by Wayne A. Hendrickson and Kurt Wüthrich. *Current Biology*: 1991/1992. Pp. 286/369. £150, \$260 (institutional); £90, \$150 (personal); £60, \$105 (student).

Protein Architecture: A Practical Approach.

By A. M. Lesk. *Oxford University Press/IRL*: 1991. Pp. 287. £32.50, \$65 (hbk); £22.50, \$45 (pbk).

A NEW industry has grown up around the explosion of protein and nucleic-acid structural information in recent years, an industry focused on the visualization and interpretation of this information. Protein crystallographers and NMR spectroscopists have ambivalent feelings about this activity. The determination of macromolecular structure is a high-risk profession with a 'Sydney or the bush' reward system: one either succeeds completely (solves the structure) or fails completely. Unlike most other areas of science, the intermediate stage of an unsuccessful project in structural biology does not usually yield publishable results. It is therefore understandable that those workers trying to determine structure regard the growing corps of structure analysts with all the affection that desert prospectors reserve for flocks of vultures.

Having often spent years struggling to solve a structure, one would like to be able to contemplate its significance oneself for a while. Yet so important has structural information become that journals and funding agencies demand public deposition of coordinates as soon as possible. And the moment that happens, the structure analysts, armed with their suites of computer programs, swoop down and extract the nuggets of insight. And woe betide the unfortunate crystallographer or NMR spectroscopist whose structure contains an error or fails to live up to the analysts' standards. It is no wonder that these exploiters of structure — most of whom are incapable of determining structures themselves — are not on many structure determiners' Christmas card lists.

So why are the feelings of crystallographers and NMR spectroscopists about this trend ambivalent? It is not a bad thing to hold structure determination to high standards, and the wealth of structural information cannot be dealt with adequately only by those who solve structures. Moreover, structure analysis has led to many computational tools of great use in structure determination.

And it has also created a demand for structure databases and methods of working with them. The two books reviewed here are designed to meet some of these needs. On the whole, they do so very well.

The *Macromolecular Structures* annuals intend to provide short summaries of data for each of the macromolecular structures determined in the previous 12 months. Each structure is given two pages, and the information provided includes the primary and some secondary references, the biological source and molecular characteristics (molecular weight, subunit composition, cofactor requirements) and a brief summary of the method of structure determination. For structures determined by X-ray crystallography, the unit-cell and space-group information are given together with the composition of the crystal mother liquor; for NMR-derived structures, the solution conditions used in the experiments are reported. In both instances, some statistics for judging the quality of the structure are tabulated, and reference is given to the status of public deposition of the atomic coordinates in the Brookhaven Protein Data Bank. Finally, a short description of the biological context of the structure is provided, together with a summary of the results and a representative illustration from the primary reference.

The 1991 volume contains 133 structure entries, and the 1992 volume is even larger. A comprehensive bibliography of all structure papers published during the year is given at the end of each volume. Although the structures are organized alphabetically, they are cross-indexed by function and by author. A book of this nature calls to mind Hilaire Belloc's verse: "The cast is large, there isn't any plot". This poses some difficulties for a reviewer. I decided that the best way to review this compendium was to live with it for some months, to keep it on my bookshelf and see how often I referred to it and what features in it I found most useful.

On average, I consulted one of the two volumes once a day, which would be more than enough to justify its purchase. I used it most often as a source of references, but I also used it frequently to find out whether coordinates for a structure had been deposited in the databank. I rarely looked at the structure-determination summary or the table of structure-quality statistics, in part because I believe that the accuracy of a structure should never be assessed by numerical data, but only by answering the question: does the structure explain the biochemical data? I found the summaries of the biological contexts and the papers to be useful, but I often wished that the summaries of the latter were

more detailed. The illustrations proved to be more decorative than useful; much of the time I wished for a different picture from the one selected.

This book is so good and so useful that one cannot resist offering suggestions to make it even better. The bibliography would be much more useful if its entries were cross-indexed by protein and author as well; as it is, I almost never consulted it. I would like more information in the structure summary and fewer crystallographic and NMR statistics, but technique enthusiasts will no doubt disagree. Each volume should also contain a complete list of all holdings in the Brookhaven Protein Data Bank, organized both alphabetically and by protein functional class. Finally, it would not be difficult to prepare two or three volumes that summarize all the structures determined before 1990. This would complete the coverage.

Protein Architecture is an addition to the IRL Practical Approach series, which includes volumes on subjects ranging from affinity chromatography to yeast. Although they are somewhat uneven, I would recommend almost any volume in the series, and this one is no exception. Arthur Lesk, who has contributed valuable computer programs and insights into structures over the past 20 years, has written a primer for scientists who want to learn how to visualize structures and extract information from them. His treatment of molecular graphics devices is clear and concise, and is followed by an excellent summary of the mathematics of molecule drawing. I am not taken with his advice on how to assess structure accuracy; he places too much value on statistical criteria. I repeat: don't believe the numbers; rather, ask whether the structure makes sense of the available biochemical data.

The second half of the book is a 60-page discussion of classes of protein structure and structure comparison. Some of this competes with the excellent new book by C. Branden and J. Tooze (*Introduction to Protein Structure*, Garland, 1991; reviewed in *Nature* **353**, 311; 1991), and I was impressed to find that the text competed well despite its lack of spectacular colour pictures. The summary of helix interactions was particularly well done and the section on protein-ligand interactions is a model of concise, clear exposition. The section on structure comparison and structural change covers material missing from Branden and Tooze's text, and is most welcome. The book ends with a complete bibliography of protein crystal structure determinations, organized according to their Brookhaven Protein Data Bank accession code. Although it is already out of date, this listing is very useful and should be adopted by *Macromolecular*

Structures, as should Lesk's classification of entries in the databank by function. The book ends with an atlas of stereo pictures of the main classes of protein fold. This sort of catalogue has been compiled by others, notably Jane Richardson, and these computer-generated drawings are not as pretty to look at as some; nevertheless, this is an excellent summary. The editors and authors of *Macromolecular Structures* and *Introduction to Protein Structure* should consider including a similar atlas, with colour to highlight secondary structure elements.

Lesk has always written well and this book is a delight to read. My only complaint is that his impish and irreverent sense of humour, which endears

him to those who know him well, is not enough in evidence here. Science books should not take themselves as seriously as they do. And scientists who determine structures should not take structure analysts as seriously as they do. But structure analysts clearly serve their purpose if they lead to books as splendidly utilitarian as these. *Protein Architecture* and the yearly issues of *Macromolecular Structures* belong on the shelves of every scientist who uses structural information. □

Gregory A. Petsko is in the Departments of Biochemistry and Chemistry and the Rosentiel Basic Medical Sciences Research Center, Brandeis University, Waltham, Massachusetts 022549110, USA.

Life and the single gene

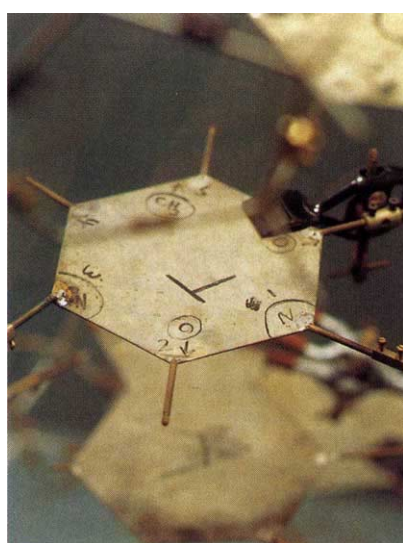
Niles Lehman

Steps Towards Life: A Perspective on Evolution. By Manfred Eigen. Oxford University Press: 1992. Pp. 173. £16.95, \$29.95.

BIOLOGISTS and physicists find themselves faced with analogous problems. Both groups of scientists are reasonably secure in their models that describe the macroscopic phenomena of their fields. And both can dissect the histories of these phenomena and even use history to understand the current structure of our surroundings. Yet there is a historical barrier that has so far proved uncrossable. For the physicist it is the origin of matter, for the biologist, the

origin of life. The biologists' 'big bang' occurred roughly four billion years ago when life on Earth coalesced from a previously abiotic environment. Darwin did not ignore this transition when he detailed the ways in which evolution could proceed through natural selection, but vacillated in his thoughts on whether natural selection could be extended back beyond the earliest living organisms.

Is the darwinian approach, at the molecular level, adequate when we try to understand the emergence of life? Manfred Eigen's answer is an elegant but forceful no. Avoiding the mathematical foundation on which his ideas are built, he carefully pieces together the logical blocks of thought that ultimately reveal the shortcomings of treating prebiotic evolution with classical 'survival-of-the-fittest' theory. The novelty is to present the origin-of-life problem no longer in a strict biological context, and instead to treat the informational enrichment of self-replicating molecules (read: Life) as a physico-chemical process tentable through a series of if-then contingencies. The traditional evolutionary approach, on which Sewall Wright relied to construct the image of populations of organisms moving through an uneven fitness landscape, and by which Jacques Monod attached the probabilistic nature of biogenesis, is rooted in stochasticity. Here, in *Steps Towards Life*, we find a de-emphasis on chance. When one concentrates on the conditions in which replication can take place, information theory allows us to perceive how the origin of life was, perhaps, inevitable. In the same way that Richard Dawkins invoked replication to explain patterns in behavioural ecology, Eigen invokes replication to explain the existence and persistence of life itself.



Models for life — reconstruction of Watson and Crick's model of DNA using the original metal plates representing the four bases. Picture taken from *Making of the Modern World* (Murray, £17.95).

The chain of arguments for the new approach is both stepwise and convoluted. We are taken through descriptions of quasi-species, error thresholds, multi-dimensional sequence space, and hypercycles in the reconstruction of the path that led to self-sustaining cell-like organisms. Each step requires precise integration into the earlier ones, and only through continual reiteration of concepts can the entire formulation avoid collapse. Eigen accomplishes this by removing the burden of explanation of many molecular biological concepts and terminology from the main body of the text and places them in both a series of separate "Vignettes" and extended sections of footnotes and glossary. Readers can therefore digest the advocacy at a variety of levels: the experienced molecular biologist will be able to read it straight through, whereas non-biologists can jump back and forth between the central argument and supplementary descriptions. Either way, the discussion is lucid, allowing a crystallization of understanding.

Steps Towards Life, which is an English translation of a work first published in 1987 in German, is unequivocally a promotion of a viewpoint that stands out in a field of study nearly as variegated as the prebiotic milieu. The quasi-species is the primary innovation, being a population of molecules of similar composition. The frequency distribution of this population is determined by each molecule's relative replication efficiency. As such, the quasi-species as a unit is subject to selective forces, in contrast to standard population-genetics models in which a wild type is predominant and all variants respond independently to selection.

It is clear how the two alternatives would fare in accommodating the origins of life. However, it is less apparent that the quasi-species concept is best — and perhaps only — applicable to acellular evolution. Eigen makes extensive use of viruses as contemporary examples of quasi-species, and moreover augments his presentation with concrete experiments done in his laboratory that are *in vitro* accounts of quasi-species evolution. Consequently, an unobservant reader may miss the brief disclaimer restricting quasi-species theory to single-gene evolution. The epigenetic interactions that play such an important role in shaping the evolution of complex organisms are not discussed, an oversight that would be of no importance if the contrast between darwinian and quasi-species evolution during the origin of life were not so strongly made. Yet in the book we are warned that darwinians are likely to defend the patterns we have come to expect in other evolutionary contexts. This is undoubtedly true.

In condensing more than 20 years of theory and experiment into a short and

engaging volume, Eigen makes the quasi-species and related concepts accessible to all scientists and thereby allows broad peer review and general consideration of his revolutionary ideas. □

Niles Lehman is in the Department of Molecular Biology, The Scripps Research Institute, 10666 North Torrey Pines Road, La Jolla, California 92037, USA.

Flower power

G. R. Moore

Pigment of the Imagination: A History of Phytochrome Research. By Linda C. Sage. Academic: 1992. Pp. 562. £66, \$99.50.

PHYTOCHROME is a blue-green pigment that acts as a plant photoreceptor protein in the mediation of many light-activated processes including seed germination, leaf formation, circadian rhythms and flowering. Such physiological processes are often sensitive to the quality, quantity or duration of light. For example, flowering of many plants is controlled by the length of day: some plants (such as lettuce and spinach) flower better under conditions of long days and short nights, whereas others (such as chrysanthemums) flower only as the days shorten (the critical factor is, in fact, the period of darkness). Horticulturists often exploit these responses, for example, by exposing chrysanthemums to a brief flash of light during the night to delay flowering until Christmas.

But precisely how the absorption of light by phytochrome results in these various activities is still not fully understood, although it has been established that the protein is photoreversible, existing in two interchangeable forms with respect to absorption, a red and a far-red form; a blue chromophore, phytochromobilin, is responsible for its activation when exposed to red light, and this photoconversion is reversible on absorption of far-red light. It is also known that there are several different species of phytochrome and that the protein determines some responses through the regulation of gene expression.

Linda Sage provides a thoroughly researched historiographical account of phytochrome research during the past 80 years, from the initial discovery that plants are able to detect daylength, through the extraction of the light-sensitive pigment, to the biophysical, biochemical and molecular genetic work of more recent years. The text is profusely illustrated with figures, tables and photographs of buildings, equipment and key researchers. The author writes in a

clear style and makes good use of footnotes to explain the science where necessary and to present important details of experiments. The result is a book that can be followed easily by nonspecialists.

In adopting a historiographical approach, Sage has had the difficult task of blending scientific facts and data with accounts of the personalities and careers of the researchers involved. The result can sometimes prove to be frustrating and compromised. For example, I was eager to learn the structure of phytochromobilin and how it attaches to its apo-protein. Although this is discussed at several points in the book, the sections are not cross-referenced well in the text, making it more difficult to grasp the current state of knowledge than if all the relevant information had been brought together in one section (there is a clear index, however). Also disappointing is the account of Garner and Allard's work. These two researchers collaborated for more than 20 years and provided the first clear demonstration of photoperiodism in plants. Yet despite the joint nature of their work it was Garner who was given much of the credit by journalists and fellow scientists for this discovery, a situation which, according to Sage, naturally displeased Allard. But exactly who did what is not made clear, and we do not get to learn Garner's thoughts on the subject. The most we are treated to is the statement that "Garner and Allard maintained outwardly cordial relations despite this tension". This sketchy treatment is a shame, because Sage has generally done a considerable amount of research, analysing scientific literature and personal records, and interviewing many people.

The publishers describe the phytochrome story as "a paradigm for the process of scientific discovery". Indeed, what Sage does very well is to remind us that scientific research builds on the work of people in a variety of disciplines, and not just those in one field. Phytochrome research provides splendid examples of the advantages of such collaboration: an understanding of the mechanisms of phytochrome action required the development of molecular biology techniques to allow the regulation of gene transcription to be studied; and an accurate characterization of plants using action spectra required the development of light filters. Too many researchers today take this collaborative nature of scientific research for granted. For this reason, the book deserves a wider audience than just photobiologists, plant physiologists and historians of science. □

G. R. Moore is in the School of Chemical Sciences, University of East Anglia, Norwich NR4 7TJ, UK.

Creative blocks: cell-cycle checkpoints and feedback controls

Andrew W. Murray

Before division, cells must ensure that they finish DNA replication, DNA repair and chromosome segregation. They do so by using feedback controls which can detect the failure to complete replication, repair or spindle assembly to arrest the progress of the cell cycle at one of three checkpoints. Failures in feedback controls can contribute to the generation of cancer.

THE division of a cell into a pair of genetically identical daughters depends on accurate chromosome replication and segregation. The fidelity of these processes depends on solving the completion problem: the need to ensure that some events in the cell cycle are completed before others begin¹. For example, eukaryotic cells must finish DNA replication before they enter mitosis, and align all their chromosomes on the mitotic spindle before they segregate them. Here I discuss two different mechanisms for solving the completion problem, citing specific examples from eukaryotes and prokaryotes, and consider how mutations that damage these mechanisms can play a critical role in the events that lead to cancer.

Figure 1 shows the division of events in the eukaryotic cell cycle into two classes: those that constitute a cell-cycle engine, and 'downstream' events that are induced by changes in the state of the engine. Downstream events are the processes involved in the actual replication and segregation of the DNA. The downstream event that marks the onset of mitosis is the assembly of the spindle, whereas the downstream event that marks the end of mitosis is chromosome segregation. The cell-cycle engine is made up of enzymes that control downstream events, and reactions that produce a cycle of coordinated changes in the activity of these enzymes. The key components of the engine are protein kinases composed of members of the p34^{cdc2} family of proteins bound to cyclins. The best characterized of these kinases is maturation promoting factor (MPF), a complex of p34^{cdc2} and cyclin B. The activation of MPF induces entry into mitosis and its inactivation induces exit from mitosis (see Box 1 for further details).

In most cell cycles, inhibiting a downstream event arrests the progress of the cell-cycle engine. For instance, inhibiting DNA replication keeps cells from entering mitosis. But in some embryos the engine continues to run even when downstream events are blocked. Fertilized frog eggs continue to show a regular alternation of MPF activation and inactivation when treated with drugs that block DNA synthesis⁶ or spindle assembly⁷. Because normal untreated embryos do not suffer from massive genetic damage, in interphase the cell-cycle engine must take longer to activate MPF than it takes to replicate DNA, so that mitosis does not occur until replication is finished. In mitosis, the engine must take longer to inactivate MPF than it takes to assemble a spindle, so that anaphase never begins until spindle assembly is complete. Thus early frog embryos rely on the steps of the cell-cycle engine being slower than the downstream events they control to solve the completion problem. Such a relative timing mechanism (Fig. 2) can only succeed in cell cycles where the rates of the downstream events and the cell-cycle engine are rigidly programmed. The cell cycles of early embryos, which divide without growing, show this rigid timing, but cells that must grow before they can divide have much more variable cell-cycle lengths. Even in cells with rigid timing, conditions that delay the downstream events relative to the engine will cause cells to undergo cell-cycle transitions before downstream events are completed and lead to death and genetic

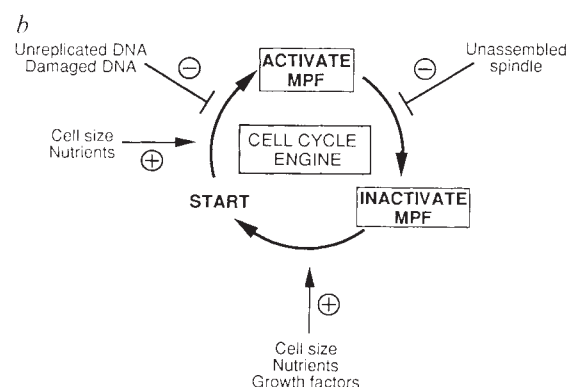
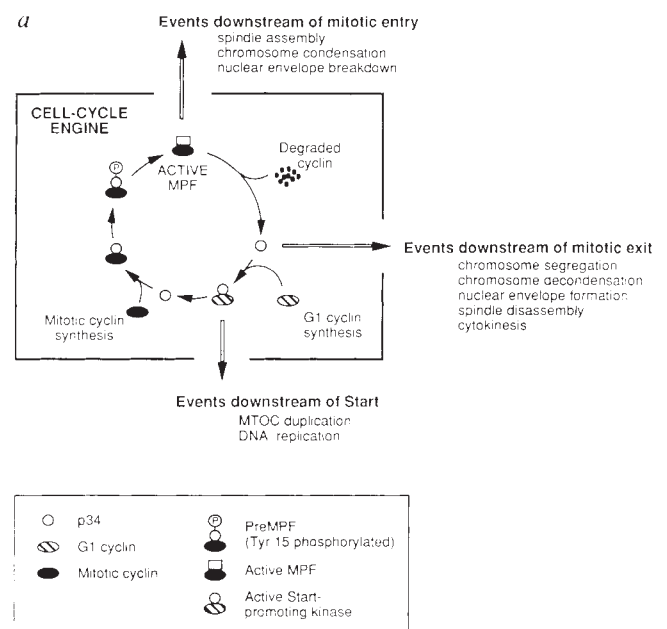


FIG. 1 The cell-cycle engine and its regulation. *a*, The cyclin and p34 components of the cell-cycle engine are represented at different points during the cell cycle, as are the downstream events that they induce. MTOC, microtubule organizing centre (the centrosome in animals and the spindle pole body in fungi). It is not known whether the degradation of G1 cyclins is regulated during the cell cycle. *b*, Feedback controls inhibit the ability of the cell-cycle engine to pass key checkpoints, while increasing cell size stimulates passage through the checkpoints. See text for further details.

damage (Fig. 2). Hartwell and Weinert suggested that this damage could be avoided if cells that failed to solve the completion problem were killed or discarded¹. In *Drosophila* embryos, nuclei that fail to complete chromosome segregation or are associated with aberrant numbers of centrosomes fall out of the layer of surface nuclei and are eliminated from development⁸.

Few cells with rigid timing rely on relative timing alone to solve the completion problem. For example, although the first few cell cycles of frog embryos are insensitive to inhibiting DNA replication, once the embryos have more than 100 nuclei, inhibiting DNA synthesis blocks entry into mitosis⁹. Unreplicated DNA is detected by a feedback control that generates a signal that prevents the activation of MPF (Fig. 2). The activation of MPF is a checkpoint for the cell-cycle engine, and feedback controls restrain the autonomous oscillation of the cell-cycle engine by regulating passage through such checkpoints. Feedback controls must have at least three components, a sensor that monitors the completion of the downstream event, a signal that this sensor generates, and a response element in the cell-cycle engine that the signal impinges on to halt or delay the engine at a checkpoint. Figure 1b shows that Start, entry into mitosis and exit from mitosis are the major checkpoints where feedback controls regulate the cell-cycle engine. Start and entry into mitosis are also used as checkpoints where cell size and the availability of nutrients and growth factors regulate the cell-cycle engine. The strength of feedback controls is highly variable. Depolymerizing spindle microtubules delays exit from mitosis for only 15 min in fertilized sea urchin eggs¹⁰, but can arrest human tissue culture cells in mitosis for many hours¹¹.

Physiology of feedback control of entry into mitosis

Finding a drug or mutation that prevents an incompleting downstream event from delaying the cell cycle demonstrates that the delay is normally due to a feedback control acting on a checkpoint. In most eukaryotic cells entry into mitosis is delayed or blocked by the presence of unreplicated or damaged DNA (reviewed in ref. 1). Caffeine treatment allows certain mammalian cells with damaged^{12,13} or unreplicated¹⁴ DNA to enter mitosis, showing that the cell-cycle arrest due to damaged or unreplicated DNA is mediated by a feedback control that can be overcome by caffeine; caffeine may be acting as a partially specific protein kinase inhibitor¹⁵.

In mammalian cells feedback controls can regulate cell growth as well as the cell-cycle engine. When synchronous cultures of human cells are treated with DNA synthesis inhibitors, the rate of protein synthesis declines and cells do not grow beyond the size of untreated cells in G2 (ref. 16), suggesting that the feedback control can regulate cell growth as well as progress of the cell-cycle engine. These cells do not suffer lethal damage and will continue to proliferate when DNA synthesis is allowed to

TABLE 1 Effect of *cdc2* phosphorylation on feedback controls in fission yeast

Genotype	Arrest by unreplicated DNA	Arrest by damaged DNA
<i>cdc25</i> overexpression	No	Partial
<i>cdc2-3w</i>	No	Yes
<i>wee1</i> ⁻	Yes	No
<i>wee1</i> ⁻ , <i>mik1</i> ⁻	No	ND
<i>cdc2-F15</i>	No	ND

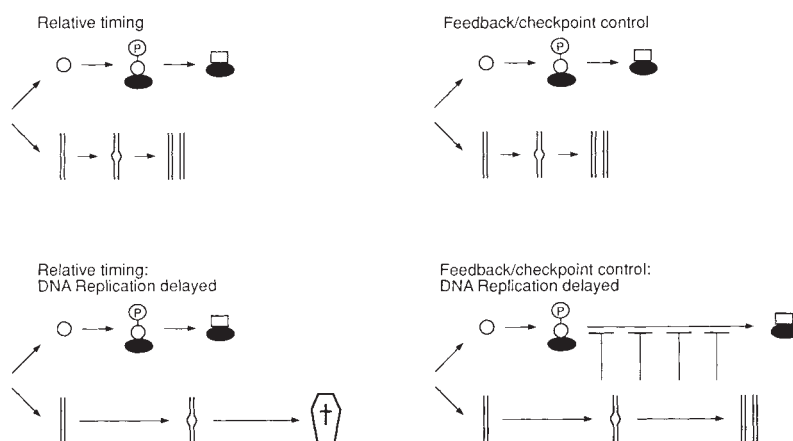
The effect of different genetic manipulations in fission yeast on the ability of unreplicated or damaged DNA to arrest the cell cycle in G2. ND, Not done. Data from refs 23, 37, 38, 41.

resume. A very different response is seen in cultured rodent cells, which continue to grow indefinitely after DNA synthesis blockade and are committed to die even if the inhibition of DNA synthesis is released. Some of the population undergo aberrant mitosis before dying, but the remainder do not, and their death may be due to apoptosis. Adding protein synthesis inhibitors, while DNA synthesis is blocked, keeps the cells from growing and spares them from death¹⁷. These results suggest that continuing to grow while arrested by feedback controls can commit cells to death by inducing programmed cell death, or by producing an increasing 'pressure' that drives the cell-cycle engine forwards, and which the feedback control cannot resist.

Genetics of feedback control of entry into mitosis

Weinert and Hartwell¹⁸ pioneered the genetic analysis of feedback controls by analysing radiation-sensitive mutations in budding yeast. They showed that the defect in three of these mutants, *rad9*, *rad17* and *rad24* was in their ability to arrest the cell cycle in response to damaged DNA, rather than in their ability to repair damaged DNA. Three novel mutants that prevent damaged DNA from arresting the cell cycle, *mec1*, *mec2* and *mec3*, have also been identified. Four of the six mutants show normal cell-cycle arrest in response to unreplicated rather than damaged DNA, but two, *mec1* and *mec2*, are also defective in their ability to arrest their cell cycle in response to unreplicated

FIG. 2 Two alternative strategies for solving the completion problem. A diagrammatic representation of the two mechanisms for coordinating the completion of DNA replication with the onset of mitosis as an example of the two different strategies for solving the completion problem. In each panel an initial event in the cell-cycle engine begins the accumulation of cyclin B in the cell-cycle engine and initiates DNA replication as a downstream event. In cells that rely on relative timing, the rates of DNA replication and MPF activation are fixed such that DNA replication is completed before MPF is activated. If DNA replication is delayed, the engine now induces cells to enter mitosis before replication is finished, resulting in death. In cells that have feedback controls, the reactions that control the activation of preMPF can be slowed down by feedback controls that are activated by unreplicated DNA, allowing cells where DNA replication has been inhibited to survive. The event that initiates the accumulation of cyclin B and DNA replication is Start in yeast cells and the exit from mitosis in early embryos. Symbols are as in Fig. 1a.



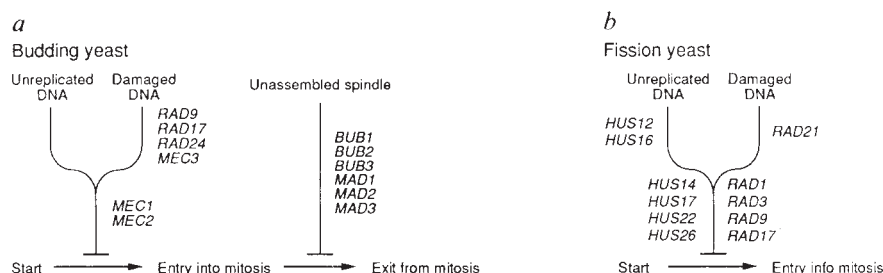


FIG. 3 Feedback controls over mitosis in the budding and fission yeasts. *a*, The genes involved in feedback control over the entry into and exit from mitosis in budding yeast. *b*, The genes involved in feedback control over the entry into mitosis in fission yeast.

DNA (T. Weinert, personal communication), suggesting that, although different sensors detect damaged and unreplicated DNA, both feedback controls converge to act on the same checkpoint in the cell-cycle engine (Fig. 3*a*). The nature of the lesions in DNA that are detected as damaged or unreplicated DNA is not known.

Deletion of the *RAD9* gene increases the rate of chromosome loss 20-fold, but is not lethal, suggesting that the feedback control that detects damaged DNA is not required in every cell cycle¹⁸. Although this control has been investigated using experimentally induced DNA damage, cells also produce molecules that are seen as damaged DNA during the course of normal metabolism, such as the single-stranded nicks and gaps produced during DNA replication. The simplest interpretation of the viability of cells lacking *RAD9* is that cells use relative timing and feedback controls as redundant ways of solving the completion problem. Even in nonembryonic cells, the relative timing of downstream events and the cell-cycle engine guarantees that in most cells the downstream events are completed before the cell-cycle engine makes its next transition. Therefore if the feedback control has been inactivated, relative timing ensures that most cells complete downstream events before the next cell-cycle transition is induced. Similarly feedback controls prevent those cells where downstream events have been delayed, either stochastically or by experimental or environmental insults,

from incurring death and genetic damage. If this analysis is correct, the feedback controls do not actively restrain the cell-cycle engine in most cycles, because relative timing ensures that on average the downstream events are completed before the cell reaches the checkpoint. In support of this argument, deletion of most^{18,19}, but not all²⁰, feedback control genes allow populations of cells to grow, albeit with increased frequencies of cell death and genetic damage. The level of redundancy between different feedback controls and relative timing is an important unresolved question.

The response to DNA damage has also been investigated in the fission yeast (Fig. 3*b*). Mutations in a number of *rad* genes allow cells with damaged or unreplicated DNA to enter mitosis^{21,22}. Other mutations allow cells with unreplicated DNA to enter mitosis, but do not affect the ability of damaged DNA to arrest the cell cycle²³. The relationship between the genes involved in budding and fission yeast feedback controls is not known.

A temperature-sensitive mutation (inactive at higher temperature) in the mammalian RCC1 (repressor of chromosome condensation) gene causes cultured hamster cells that are shifted-up to the nonpermissive temperature during S phase to cease DNA replication and enter mitosis prematurely. Cells shifted up early in G1 do not synthesize any DNA after shift-up and do not enter mitosis²⁴. This mutation clearly affects both

BOX 1 A cell cycle refresher

THE past five years have seen a prodigious advance in our understanding of the cell cycle, which is very briefly reviewed here. The key component proteins of the cell-cycle engine are members of the p34 family and cyclins. The protein p34^{cdc2} was identified as the product of two cell division cycle genes, *cdc2* of the fission yeast, *Schizosaccharomyces pombe*, and *CDC28* of the budding yeast, *Saccharomyces cerevisiae*. Cyclins were discovered by studying patterns of protein synthesis in fertilized sea urchin eggs. The stability of cyclin proteins varies throughout the cell cycle, whereas p34^{cdc2} is stable. Cyclins are divided into a number of classes based on sequence similarity and physiological function: mitotic cyclins (cyclin B), S-phase cyclins (cyclin A) and G1 cyclins (called Cln in yeast and cyclins C, D, E and F in vertebrates).

Figure 1*a* shows the cell cycle of a yeast cell. In G1, the beginning of the cell cycle, p34^{cdc2} lacks any associated cyclins and has no protein kinase activity. As long as cells are adequately supplied with nutrients and sufficiently large, they accumulate G1 cyclins that bind to p34^{cdc2} and activate its protein kinase activity, thereby inducing cells to pass through Start, the first transition in the cell cycle. The Start-promoting activity of p34^{cdc2} induces cells to replicate their DNA (entering S phase) and their microtubule organizing centre (MTOC). In yeasts the molecular regulation of Start is fairly well understood, whereas our understanding of how mammalian cells pass the restriction point (the equivalent of Start) is much more fragmentary. In mammalian cells and budding yeast the restriction point/Start is the principal point in the cell cycle where proliferation is regulated by nutrient availability and growth factors (Fig. 1*b*).

After passing Start, cells begin to synthesize cyclin B which accumulates and binds to p34^{cdc2}. This complex is inactive because the binding of cyclin B to p34^{cdc2} induces the phosphorylation of tyrosine 15 of p34^{cdc2}, which inhibits its protein kinase activity. The inactive complex, often referred to as preMPF, is also phosphorylated on threonine 160. This phosphorylation is required for MPF activity but is not sufficient to overcome the inhibitory effect of tyrosine phosphorylation. The subsequent

removal of the tyrosine phosphate from p34^{cdc2} during late G2 activates MPF and leads to the induction of mitosis (M phase). These post-translational reactions play a major role in regulating the timing of MPF activation and are subject to complex physiological regulation. The tyrosine kinase was initially identified as the product of the fission yeast *wee1* gene, which acts as an inhibitor of entry into mitosis, whereas the phosphatase is the product of the fission yeast *cdc25* gene which acts to promote entry into mitosis. Active MPF seems to activate the tyrosine phosphatase and inhibit the tyrosine kinase that modifies p34^{cdc2}, leading to an explosive activation of MPF that drives cells rapidly and irreversibly into mitosis. MPF induces chromosome condensation, nuclear envelope breakdown and assembly of the mitotic spindle. MPF induces nuclear envelope breakdown directly by phosphorylating the nuclear lamins, but it is not known whether chromosome condensation and spindle assembly are induced directly by MPF or indirectly by other enzymes activated by MPF.

Once MPF is active it induces the activation of enzymes that conjugate ubiquitin to cyclin and thereby target cyclin for degradation by a well characterized system that proteolyse poly-ubiquitinated proteins. The degradation of cyclin destroys the kinase activity of MPF and induces the cell cycle to progress into the next interphase and allows cyclin to become stable again. The degradation of cyclin and the inactivation of MPF induce chromosome segregation, chromosome decondensation, nuclear reformation and cytokinesis.

Somatic and embryonic cell cycles differ mainly in the regulation of DNA synthesis and microtubule organizing centre duplication. In embryonic cell cycles DNA and microtubule organizing centre replication are downstream events induced by the inactivation of MPF rather than by passage through Start. Both of these events occur as soon as nuclear envelopes have reformed after the inactivation of MPF, and there is no evidence for any equivalent of Start in the embryonic cell cycle. For further details and full references see refs 2–5.

feedback controls and cell-cycle timing. The RCC1 protein is found in chromatin²⁵ and can stimulate guanine nucleotide exchange on a small G protein called *ran*, whose function is unknown²⁶. Perhaps the *ran* protein is a component required for the stability of replication forks and the feedback control works by detecting replication forks. Once the RCC1 protein has been inactivated, the *ran* protein accumulates in the inactive GDP-bound form causing replication forks to disappear. As a result, the feedback control machinery reports that DNA replication is complete, and allows the cell to enter mitosis.

A fission yeast mutation in a gene related to RCC1 also affects entry into mitosis²⁷. The mutant accumulates cells in mitosis when incubated at the nonpermissive temperature, and the mutation lies in the *pim1* gene which encodes a protein with 30% sequence identity to RCC1. Mutant cells can enter mitosis even when they have not passed Start, and this phenotype can be suppressed by overexpressing a small G protein²⁷. This protein's closest mammalian relative is not the *ran* protein but another small G protein called TC4 (ref. 28), whose function is also unknown. Relatives of RCC1 have also been identified in budding yeast^{29,30} and *Drosophila*³¹, and both genes can complement the mammalian RCC1 mutation³². But mutants in the budding yeast homologue of RCC1 are defective in splicing³⁰ and the response to mating pheromone²⁹, rather than in the ability to respond normally to unreplicated DNA. It remains to be seen whether the yeast and *Drosophila* proteins are true homologues of RCC1, or members of a superfamily of guanine nucleotide exchange proteins with overlapping but not identical functions.

In the fungus *Aspergillus nidulans*, the temperature-sensitive *bimE* mutant allows cells with unreplicated DNA to enter mitosis³³. The *bimE* cells enter mitosis from any stage of the cell cycle when they are shifted to the nonpermissive temperature. Mutations in *bimE* interact with mutations in *nimA*, which encodes a protein kinase whose activity is required for entry into mitosis³⁴ and which is activated independently of MPF (ref. 35). The *nimA*, *bimE* double mutants can still enter mitosis, suggesting that BimE may act as an inhibitor of entry into mitosis that must be overcome by the action of NimA (refs 33, 34). Perhaps NimA is a responder element that inactivates the BimE protein to signal to the cell-cycle engine that replication is complete (Fig. 4), although NimA activity is also required for some of the structural events of mitosis³⁴.

Mutations in components of the cell-cycle engine can affect the operation of feedback controls (Table 1). The most striking examples are mutations in the genes that control the phosphorylation of tyrosine 15 in p34 (see Box 1). This modification inhibits kinase activity and prevents entry into mitosis³⁶. Strong overexpression of the tyrosine phosphatase (Cdc25), dominant activating alleles of *cdc2* (*cdc2-3w*), mutation of the tyrosine in p34 to phenylalanine, or inactivation of the tyrosine kinases (Wee1 and Mik1) all accelerate entry into mitosis and abolish the ability of unreplicated DNA to block entry into mitosis in fission yeast^{23,37,38} and frog embryos^{39,40}. Modifications to p34 also affect the response of fission yeast to DNA damage. *wee1* cells cannot arrest their cell cycle in response to low doses of irradiation, but arrest in response to high doses of irradiation or inhibition of DNA synthesis with hydroxyurea⁴¹.

The difference between the response of the fission yeast *rad1*

and *cdc2-3w* mutations to unreplicated DNA illustrates the difference between lesions in feedback control components and those in the cell-cycle engine response elements that constitute a checkpoint. The *rad1* mutant acts as a feedback control mutant: in the absence of unreplicated DNA the mutant and wild-type cells enter mitosis at the same time, but in *rad1* cells the presence of unreplicated DNA does not delay entry into mitosis. In contrast, the *cdc2-3w* mutant acts as if an intact feedback control is trying to restrain a runaway cell-cycle engine: in the absence of unreplicated DNA the mutant enters mitosis earlier than wild-type cells, but even in *cdc2-3w* the presence of unreplicated DNA does briefly delay entry into mitosis²¹.

Not all cells rely on inhibiting tyrosine dephosphorylation of p34 to respond to unreplicated DNA. In budding yeast, p34 mutations that prevent tyrosine phosphorylation have no effect on the ability of cells to detect and respond to unreplicated or damaged DNA (ref. 42). One reservation about this result is that in normal budding yeast cycles, spindle assembly⁴³ and the activation of MPF can start before DNA replication is complete^{42,44}. But two observations suggest that budding yeast still has a discrete cell-cycle transition late in the cell cycle that corresponds to entry into mitosis in other organisms: the level of MPF activity in cells arrested by inhibitors of microtubule assembly is higher than in cells arrested by inhibitors of DNA synthesis, and the phosphorylation of p34 differs between cells arrested with inhibitors of DNA synthesis and cells arrested with inhibitors of spindle assembly^{42,44}. The biochemical nature of the postreplication transition in the budding yeast cell-cycle engine remains unknown.

Feedback control over exit from mitosis

Cells exert feedback control over exit from mitosis in response to the completion of spindle assembly. Two groups have isolated budding yeast mutants, called *bub* (budding uninhibited by benomyl)¹⁹ or *mad* (mitotic arrest deficient)²⁰, which abolish this feedback control (Fig. 3a). In both cases, the mutations have no obvious effect on the number or stability of microtubules, but are unable to maintain high levels of MPF activity in the presence of microtubule depolymerizing agents. This feedback control probably acts by preventing the activation of the machinery that degrades the mitotic cyclins and leads to the inactivation of MPF (ref. 45).

All three of the *BUB* genes and one *MAD* gene have been cloned. *BUB1* resembles serine/threonine protein kinases (M. A. Hoyt, personal communication), but *BUB2* and *BUB3* lack sequence similarity to known proteins¹⁹. *MAD2* is homologous to the α -subunit of a prenyl transferase⁴⁶. Prenyl groups are attached to the carboxy termini of a number of proteins to anchor them to membranes. Many of the proteins thus modified are GTP-binding proteins, including the *ras* protein, other small G proteins and the γ -subunits of trimeric G proteins (reviewed in ref. 47). The involvement of *MAD2* and RCC1 in two different feedback controls suggests that G proteins may be widely involved in feedback controls.

Several lines of evidence implicate microtubule organizing centres in feedback controls. First, mammalian cells that have had their centrosomes surgically removed during interphase fail to enter mitosis⁴⁸. Because cells treated with microtubule polymerization inhibitors enter mitosis normally, it seems that a

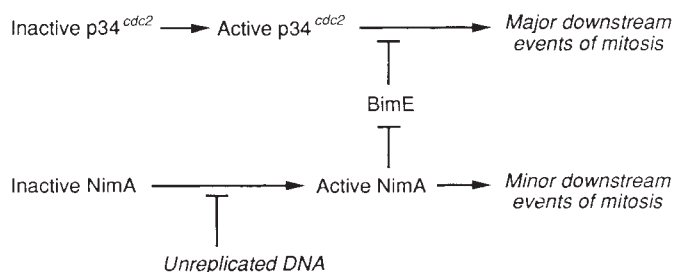


FIG. 4 Control over mitotic entry in *Aspergillus*. A speculative scheme for the dual role of p34 and NimA in inducing mitosis in *Aspergillus*. The p34-associated kinase is responsible for inducing most of the downstream events of mitosis, but some events also require the NimA-associated kinase for their successful completion. The NimA-associated kinase may be regulated by completion of DNA replication, and function to overcome the inhibitory effect of BimE.

feedback control monitors the presence of the centrosomes rather than detecting the microtubules that they nucleate. Second, *Aspergillus* cells that lack γ -tubulin, a component of the microtubule organizing centre, can leave mitosis normally even though they lack microtubules, suggesting that the feedback control over mitosis requires an intact microtubule organizing centre⁴⁹. Finally, praying mantis spermatocytes are able to detect chromosomes that are located too close to the spindle poles in mitosis and activate a feedback control that prevents exit from mitosis (reviewed in ref. 50).

Feedback control and fertilization

Unfertilized eggs are arrested at a variety of stages in the cell cycle, and this arrest is relieved by fertilization. The cell-cycle arrest can be thought of as a feedback control that detects the failure to complete an external event, sperm-egg fusion, rather than an internal one. Unfertilized frog eggs are arrested in metaphase of meiosis II, and this arrest is mediated by an activity known as cytosolic factor (CSF)⁵¹. CSF is intimately related to the product of the *c-mos* proto-oncogene⁵², a protein kinase that in some way blocks the ability of the MPF to induce the degradation of cyclin and thereby induce its own inactivation⁵³. Fertilization induces an increase in intracellular calcium concentration that overcomes the ability of *c-mos* to prevent cyclin degradation, and the subsequent inactivation of MPF seems to induce the proteolysis of *c-mos* (refs 54, 55). This is one case where completely unambiguous evidence links a calcium transient to the release of a feedback control on the cell cycle. The role of calcium transients in controlling entry into and exit from mitosis in somatic cells is much less clear (reviewed in ref. 56).

Feedback controls in prokaryotic cells

Eukaryotes must ensure that cell division does not occur until DNA replication and repair have been completed. Because neither of these processes can occur after chromosome condensation, eukaryotes must ensure that replication and repair are complete before entering mitosis. Because chromosome segregation and cell division are both initiated by the inactivation of MPF, eukaryotes cannot monitor the success of chromosome segregation but must resort to regulating the inactivation of MPF in response to the assembly of the chromosome segregation machinery. The simpler organization of prokaryotes makes it possible to couple cell division directly to the completion of replication, the repair of DNA damage and chromosome segregation. Bacteria divide by forming a septum and the key inducer of septation is the FtsZ protein, which becomes localized at the septation site just before septation begins⁵⁷. (See refs 58, 59 for reviews of the bacterial cell cycle.)

In rapidly growing bacteria the coordination of DNA replication and cell division is complicated because the cell cycle is actually shorter than the time it takes to replicate the chromo-

some. To overcome this problem, cells initiate a round of chromosome replication in one cell cycle and complete it in the next. Figure 5a shows that this strategy means that rapidly growing cells possess at least one set of active replication forks at all points during the cell cycle. Thus procaryotic cells initiate and complete cell division during active DNA replication. At first sight, it seems that procaryotes, unlike eukaryotes, use a positive control to coordinate cell division and DNA replication: the act of terminating a round of replication induces cell division. But Begg and Donachie⁶⁰ showed that septation is inhibited by nucleoids near the presumptive septation site and suggested that termination induces cell division indirectly by permitting chromosome segregation. Before termination the unsegregated chromosomes lie in the centre of the cell near the site where septation will occur. As a result, septation cannot occur until the two daughter nucleoids have separated from each other and moved away from the presumptive septation site towards the ends of the cell. The presence of nucleoids near the presumptive septation site represents an uncompleted downstream event that activates a feedback inhibition of cell division, so that the logic of division control in prokaryotes and eukaryotes is fundamentally similar (Fig. 5b).

In prokaryotes, like eukaryotes, DNA damage blocks cell division by activating a feedback control system that prevents cell division (Fig. 5b). This control is the part of the 'SOS' system that bacterial cells use to respond to damaged DNA (reviewed in ref. 61). Damaged DNA induces the RecA protein to act as a cofactor in the proteolysis of the LexA protein, allowing numerous genes whose expression is involved in the cellular response to DNA damage to escape from repression by LexA. One of these genes is *sfiA* (also called *sulA*), which encodes a protein that inhibits the activity of the FtsZ protein⁶².

Feedback controls and cancer

Lesions in feedback controls may play a critical role in the induction and progression of cancer. Epidemiological⁶³ and molecular⁶⁴ studies show that the induction of cancer in mammals requires the accumulation of several independent mutations. Mutations that affect growth control and cellular behaviours involved in metastasis are not the only changes that can affect the progression of cancer. Any change that increases the chance of genetic damage, will increase the rate at which cells accumulate mutations that make them malignant. This is especially true for recessive mutations that inactivate genes that restrain cell proliferation, and which cannot be expressed until wild-type copies of the gene have been removed by mitotic recombination or chromosome loss. Budding yeast mutants that inactivate the feedback controls that detect damaged DNA or errors in spindle assembly have precisely this effect: they both elevate the frequency of chromosome loss about 30-fold (refs 18, 20).

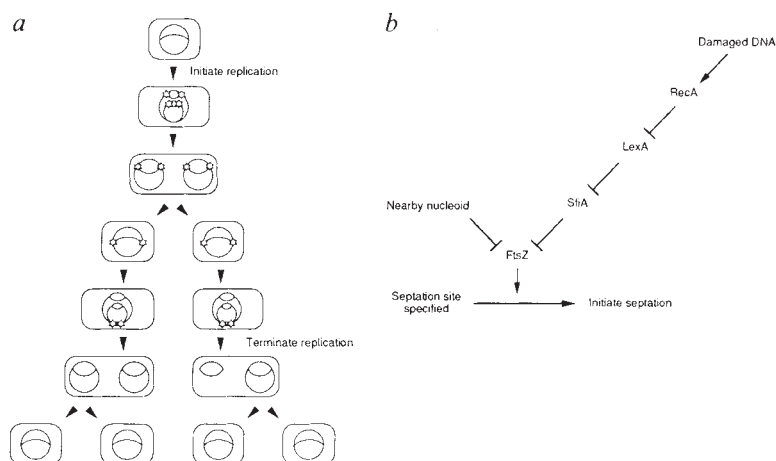


FIG. 5 Feedback controls in *Escherichia coli*. *a*, The pattern of DNA replication in rapidly growing bacteria. A replication origin that fires in one cell cycle gives rise to replication forks (indicated by stars) that complete replication in the subsequent cell cycle. Each termination event is followed by chromosome segregation. *b*, Feedback control of cell division. The ability of the FtsZ protein to initiate septation is inhibited by the SfiA protein, whose expression is induced by DNA damage, and possibly by nearby nucleoids. See text for further details.

In mammalian cells DNA damage caused by mild X-ray irradiation can block passage through two cell-cycle checkpoints, the restriction point^{65,66} and the entry into mitosis. In addition, moderate X-ray irradiation during S phase leads to a rapid block in the firing of new replication origins that allows DNA damage to be repaired before S phase continues⁶⁷. Ataxia telangiectasia (AT) is a human disease caused by a recessive mutation that inactivates the feedback control that regulates DNA replication in irradiated cells (reviewed in ref. 68). Patients suffer from damage to blood vessels in the skin, immune system deficiencies, problems in muscle coordination, and greatly increased incidence of lymphoid cancers. Cell lines from AT patients are hypersensitive to killing by X-rays⁶⁹, and they continue to initiate new replication origins after irradiation⁷⁰. In untransformed AT fibroblasts, irradiation also fails to cause arrest in G2 (refs 71, 72), or to prevent the passage of cells through the restriction point⁷³, suggesting that AT cells are deficient in all responses to damaged DNA.

The mammalian p53 protein plays a crucial role in human cancers. The protein is believed to act as brake on the cell cycle by activating transcription of genes whose products inhibit progress through the restriction point, and mutations in p53 are the most commonly found mutations in human cancer (reviewed in ref. 74). Recent evidence suggests that p53 is part of the feedback control whereby DNA damage arrests cells at the restriction point. Mild ultraviolet or X-ray irradiation^{75,76} elevates the level of p53 protein by increasing its half-life⁷⁵. Tumour cell lines that lack p53 or express dominant negative mutants of the protein do not show X-ray-induced G1 arrest, although they still arrest in G2 in response to irradiation⁷⁶. Li-Fraumeni syndrome is a genetic predisposition to develop cancer⁷⁷ that is associated with point mutations in p53 (refs 78, 79). Fibroblasts derived from Li-Fraumeni patients show normal growth properties, but have a high frequency of spontaneous chromosomal abnormalities and when passaged in culture eventually give rise

to morphologically transformed and immortalized cells⁸⁰. Taken together, these results suggest that a major function of p53 protein is in the feedback control that keeps cells from entering S phase while they contain damaged DNA. A number of the other recessive mutations that are commonly found in human tumours may also inactivate functions involved in feedback controls, rather than in the regulation of cell proliferation.

Understanding feedback controls should lead to better treatments for cancer. Most chemotherapeutic agents block DNA replication, damage DNA or interfere with chromosome segregation. The lethal event for cells treated with inhibitors of DNA metabolism is entry into mitosis, whereas that for cells treated with inhibitors of spindle function is the onset of anaphase. Byfield *et al.*⁸¹ suggested that treating cells with inhibitors of feedback controls could increase the potency of drugs that damaged DNA, and caffeine⁸² and related compounds⁸³ do increase the killing of tumour cells by DNA alkylating agents. Because tumour cells may have been selected to have weaker feedback controls, feedback control inhibitors might increase the selectivity, as well as the cytotoxicity, of chemotherapeutic drugs. Examination of rodent cell lines shows that feedback controls do indeed become weaker as cells become more transformed. Primary cultures show strong feedback controls, established cell lines have somewhat weaker ones, and those of transformed lines are weaker still⁸⁴. With specific inhibitors of feedback controls it might be possible to weaken further the feedback controls of tumour cells but have little effect on those of normal cells, leading to more selective killing of tumour cells. These considerations suggest that studies of feedback controls, and an extensive search for inhibitors of feedback controls may have important medical consequences, as well as increasing our knowledge of how cells safeguard their genetic dowries. □

Andrew W. Murray is the Department of Physiology, University of California at San Francisco, Box 0444, San Francisco, California 94143-0444, USA.

- Hartwell, L. H. & Weinert, T. A. *Science* **246**, 629-634 (1989).
- Murray, A. W. & Kirschner, M. W. *Science* **246**, 614-621 (1989).
- Nurse, P. *Nature* **344**, 503-508 (1990).
- Draetta, G. *Trends biochem. Sci.* **15**, 378-383 (1990).
- Cycle, T. C. *Cold Spring Harb. Symp. quant. Biol.* **66**, 782 (1991).
- Kimelman, D., Kirschner, M. & Scherson, T. *Cell* **48**, 399-407 (1987).
- Gerhart, J., Wu, M. & Kirschner, M. *J. Cell Biol.* **98**, 1247-1255 (1984).
- Sullivan, W., Minden, J. S. & Alberts, B. M. *Development* **110**, 311-323 (1990).
- Dasso, M. & Newport, J. W. *Cell* **61**, 811-823 (1990).
- Sluder, G., Miller, F. J. & Spangian, K. *J. exp. Zool.* **238**, 325-336 (1986).
- Zeive, G. W., Turnbull, D., Mullins, J. M. & McIntosh, J. R. *Exp. Cell Res.* **126**, 397-406 (1980).
- Boynton, A. L., Evans, T. C. & Crouse, D. A. *Radiation Res.* **60**, 80-97 (1974).
- Lau, C. C. & Pardee, A. B. *Proc. natn. Acad. Sci. U.S.A.* **79**, 2942-2946 (1982).
- Schlegel, R. & Pardee, A. B. *Proc. natn. Acad. Sci. U.S.A.* **84**, 9025-9029 (1987).
- Schlegel, R., Belinsky, G. S. & Harris, M. O. *Cell Growth Differ.* **1**, 171-178 (1990).
- Schimke, R. T., Kung, A. L., Rush, D. F. & Sherwood, S. W. *Cold Spring Harb. Symp. quant. Biol.* **66**, 417-425 (1991).
- Kung, A. L., Zetterburg, A., Sherwood, S. W. & Schimke, R. T. *Cancer Res.* **50**, 7307-7317 (1990).
- Weinert, T. A. & Hartwell, L. H. *Science* **241**, 317-322 (1988).
- Hoyt, M. A., Trotis, L. & Roberts, B. T. *Cell* **66**, 507-517 (1991).
- Li, R. & Murray, A. W. *Cell* **66**, 519-531 (1991).
- Rowley, R., Subramani, S. & Young, P. G. *EMBO J.* **11**, 1335-1342 (1992).
- Al-Khodairy, F. & Carr, A. M. *EMBO J.* **11**, 1343-1350 (1992).
- Enoch, T., Gould, K. L. & Nurse, P. *Cold Spring Harb. Symp. quant. Biol.* **56**, 409-416 (1991).
- Nishimoto, T., Ellen, E. & Basilico, C. *Cell* **15**, 475-483 (1978).
- Ohtsubo, M., Okazaki, H. & Nishimoto, T. *J. Cell Biol.* **109**, 1389-1397 (1989).
- Bischoff, F. R. & Ponstingl, H. *Nature* **354**, 80-82 (1991).
- Matsumoto, T. & Beach, D. *Cell* **66**, 347-360 (1991).
- Drivas, G. T., Shih, A., Coutavas, E., Rush, M. G. & D'Eustachio, P. *Molec. cell. Biol.* **10**, 1793-1798 (1990).
- Clark, K. L. & Sprague, G. F. *Molec. cell. Biol.* **9**, 2682-2694 (1989).
- Aebi, M., Clark, M. W., Vijayraghavan, U. & Abelson, J. *Molec. Gen. Genet.* **224**, 72-80 (1990).
- Frasch, M. *EMBO J.* **10**, 1225-1236 (1991).
- Ohtsubo, M., *et al.* *EMBO J.* **10**, 1265-1273 (1991).
- Osmani, S. A., Engle, D. B., Doonan, J. H. & Morris, N. R. *Cell* **52**, 241-251 (1988).
- Osmani, A. H., O'Donnell, K., Pu, R. T. & Osmani, S. A. *EMBO J.* **10**, (1991).
- Osmani, A. H., McGuire, S. L. & Osmani, S. A. *Cell* (1991).
- Gould, K. L. & Nurse, P. *Nature* **342**, 39-45 (1989).
- Enoch, T. & Nurse, P. *Cell* **60**, 665-673 (1990).
- Lundgren, K. *et al.* *Cell* **64**, 1111-1122 (1991).
- Kumagai, A. & Dunphy, W. G. *Cell* **64**, 903-914 (1991).
- Smythe, C. & Newport, J. W. *Cell* **68**, 787-797 (1992).
- Rowley, R., Hudson, J. & Young, P. G. *EMBO J.* (1992).
- Sorger, P. K. & Murray, A. W. *Nature* **355**, 365-368 (1992).
- Byers, B. in *The Molecular Biology of the Yeast Saccharomyces. Life Cycle and Inheritance* (eds Strathern, J. N., Jones, E. W. & Broach, J. R.) 59-96 (Cold Spring Harbor Laboratory Press, New York, 1981).
- Amon, A., Surana, U., Muroff, I. & Nasmyth, K. *Nature* **355**, 368-371 (1992).
- Glutzer, M., Murray, A. W. & Kirschner, M. W. *Nature* **349**, 132-138 (1991).
- Boguski, M., Murray, A. W. & Powers, S. *New Biol.* (in the press).
- Glomset, J., Gelb, M. & Farnsworth, C. *Curr. Opin. Lipidol.* **2**, 118-124 (1992).
- Maniotis, A. & Schliwa, M. *Cell* **67**, 495-504 (1991).
- Oakley, B. R., Oakley, C. E., Yoon, Y. & Jung, M. K. *Cell* **61**, 1289-1301 (1990).
- Nicklas, R. B. & Arana, P. *J. Cell Sci.* (in the press).
- Masui, Y. & Markert, C. L. *J. exp. Zool.* **277**, 129-145 (1971).
- Sagata, N., Watanabe, N., Vande Woude, G. F. & Ikawa, Y. *Nature* **342**, 512-518 (1989).
- Murray, A. W., Solomon, M. J. & Kirschner, M. W. *Nature* **339**, 280-286 (1989).
- Lorca, T. *et al.* *EMBO J.* **10**, 2087-2093 (1991).
- Watanabe, N., Hunt, T., Ikawa, Y. & Sagata, N. *Nature* **352**, 247-248 (1991).
- Hepler, P. K. *J. Cell Biol.* **109**, 2567-2573 (1989).
- Bi, E. & Lutkenhaus, J. *Nature* **354**, 161-164 (1991).
- de Boer, P. A. J., Cook, W. R. & Rothfield, L. I. *A. Rev. Genet.* **24**, 249-274 (1990).
- Donachie, W. D. & Begg, K. *Res. Microbiol.* **141**, 64-75 (1990).
- Begg, K. J. & Donachie, W. D. *New Biol.* **3**, 475-486 (1991).
- Walker, G. C. in *Escherichia coli and Salmonella typhimurium: Cellular and Molecular Biology* (ed. Neidhart, F. C.) 1346-1357 (Am. Soc. Microbiol., Washington DC, 1987).
- Lutkenhaus, J. *Trends Genet.* **6**, 22-25 (1990).
- Peto, R., Roe, F. J., Lee, P. N., Levy, L. & Clack, J. *Br. J. Cancer* **32**, 411-426 (1975).
- Frearon, E. R. & Vogelstein, B. *Cell* **61**, 759-767 (1990).
- Little, J. B. *Nature* **218**, 1064-1065 (1968).
- Nagasawa, H., Robertson, J., Arundel, C. S. & Little, J. B. *Radiation Res.* **97**, 537-545 (1984).
- Painter, R. B. & Young, B. R. *Radiation Res.* **64**, 648-656 (1975).
- Cohen, M. M. & Levy, H. P. *Adv. hum. Genet.* **18**, 43-149 (1989).
- Taylor, A. M. R. *et al.* *Nature* **258**, 427-429 (1975).
- Painter, R. B. & Young, B. R. *Proc. natn. Acad. Sci. U.S.A.* **77**, 7315-7317 (1980).
- Zampetti-Bosseler, F. & Scott, D. *Int. J. Radiat. Biol.* **39**, 547-558 (1981).
- Rudolph, N. S. & Latt, S. A. *Mutation Res.* **211**, 31-41 (1989).
- Nagasawa, H., Latt, S. A., Lalande, M. E. & Little, J. B. *Mutation Res.* **148**, 71-82 (1985).
- Lane, D. P. *Nature* **358**, 15-16 (1992).
- Maltzman, W. & Czyzyk, L. *Molec. cell. Biol.* **4**, 1689-1694 (1984).
- Kastan, M. B., Onyekwere, O., Sidransky, D., Vogelstein, B. & Craig, R. W. *Cancer Res.* **51**, 6304-6311 (1991).
- Li, F. P. & Fraumeni, J. F. *J. Am. med. Assoc.* **247**, 2692-2694 (1982).
- Malkin, D. *et al.* *Science* **250**, 1233-1238 (1990).
- Srivasta, S., Zou, Z., Pirollo, K., Blattner, W. & Chang, E. H. *Nature* **348**, 747-749 (1990).
- Bischoff, F. Z. *et al.* *Cancer Res.* **50**, 7979-7984 (1990).
- Byfield, J. E. *et al.* *Br. J. Cancer* **43**, 669-683 (1981).
- Gaudin, D. & Yielding, K. L. *Proc. Soc. exp. Biol. Med.* **131**, 1413-1416 (1969).
- Fingert, H. J. *et al.* *Cancer Res.* **48**, 4375-4381 (1989).
- Kung, A. L., Sherwood, S. W. & Schimke, R. T. *Proc. natn. Acad. Sci. U.S.A.* **87**, 9553-9557 (1990).

ACKNOWLEDGEMENTS. I thank my colleagues in the cell-cycle field for many stimulating discussions and for communicating their results before publication and especially L. Hartwell, E. Weinert, M. Kirschner, A. Hoyt, P. Sorger and T. Hyman for their critical reading of the manuscript. I am a Lucille P. Markey scholar and a David and Lucile Packard fellow. Work in my laboratory is supported by NIH, the Markey Charitable Trust, the Packard foundation and the March of Dimes.

Carbon isotope evidence for the stepwise oxidation of the Proterozoic environment

David J. Des Marais*, Harald Strauss†, Roger E. Summons‡ & J. M. Hayes§

* NASA, Ames Research Center, Moffett Field, California 94035, USA

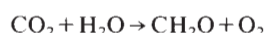
† Institut für Geologie, Ruhr-Universität Bochum, Postfach 102148, 4630 Bochum 1, Germany

‡ Bureau of Mineral Resources, GPO Box 378, Canberra, Australian Capital Territories 2605, Australia

§ Biogeochemical Laboratories, Departments of Geological Sciences and of Chemistry, Geology Building, Indiana University, Bloomington, Indiana 47405, USA

The oxidation of the Earth's crust and the increase in atmospheric oxygen early in Earth history have been linked to the accumulation of reduced carbon in sedimentary rocks. Trends in the carbon isotope composition of sedimentary organic carbon and carbonate show that during the Proterozoic aeon (2.5–0.54 Gyr ago) the organic carbon reservoir grew in size, relative to the carbonate reservoir. This increase, and the concomitant release of oxidizing power in the environment, occurred mostly during episodes of global rifting and orogeny.

PERHAPS life's most extensive alteration of its environment arises from the splitting of the water molecule during oxygenic photosynthesis to produce molecular oxygen (O_2) and reducing power for the synthesis of organic matter



This environmental alteration became extensive when part of the O_2 escaped back-reaction with organic matter (CH_2O in the equation above), because some of the organic matter was buried in sediments and thus removed. Therefore O_2 , sulphate (SO_4^{2-}) and ferric iron (Fe^{3+}) accumulated in the crust and atmosphere largely because organic matter and biogenic sulphide accumulated in sediments^{1,2}. Even biogenic sulphide owes its existence largely to sedimented organic matter, which fuels bacterial reduction of sulphate. Oxidation of the Earth's surface environment was thus promoted not by biological processes alone, but by their interaction with geological processes such as sedimentation.

The history of oxidation during the Proterozoic aeon can be traced by quantifying the growth of the crustal organic carbon reservoir^{2,3}. Each mole of organic carbon represents a mole of O_2 that either still exists or has reacted with some oxidizable substance. Most important among the latter are Fe^{2+} and reduced sulphur. The delivery of these to the surface environment is also controlled by geological processes.

With these factors in mind, we present here the results of a coordinated study of the Earth's biogeochemical and tectonic records. We have used new data and have taken into account processes occurring since the initial deposition of the organic material. Probable rates of deposition and recycling of sediments have been considered. We find that oxidation of the environment has proceeded very unevenly and that the main transitions are correlated with geological rather than biological events.

Mass balances within the carbon cycle

Stable isotope analyses of carbonate and organic matter offer the most effective approach for tracing the growth of the crustal reservoir of reduced carbon. Natural abundances of ^{13}C in marine sedimentary carbonates ($\delta_{carb} = (^{13}C/^{12}C)_{carb}/(^{13}C/^{12}C)_{standard} - 1$) and in organic matter (δ_{org}) are controlled by (1) equilibrium isotope effects among inorganic carbon species, (2) isotopic fractionations associated with the biosynthesis and reworking of organic matter, and (3) the relative rates of immobilization and burial of carbonate and organic carbon in sediments. The isotopic difference between buried carbonate carbon and organic carbon, $\Delta_c = \delta_{carb} - \delta_{org}$, is therefore an indicator of processes within the global carbon cycle.

Isotope effects associated with the enzymatic fixation of carbon and with mass transport of CO_2 are responsible for most of Δ_c . The overall fractionation accompanying photosynthesis can be attenuated by limitations in the supply of CO_2 (refs 4, 5). The amounts of carbonate and organic carbon buried depend on global rates of erosion and sedimentation⁶ and on recycling processes at the sea floor⁷. The operation of the carbon cycle can be monitored through an isotopic mass balance

$$\delta_{in} = f_{carb}\delta_{carb} + f_{org}\delta_{org} \quad (1)$$

where δ_{in} represents the isotopic composition of carbon entering the global surface environment comprised of the atmosphere, hydrosphere and biosphere, and the right side of the equation represents the weighted-average isotopic composition of carbon being buried in sediments, f_{carb} and f_{org} being the fractions of carbon buried in inorganic and organic form ($f_{carb} = 1 - f_{org}$). Carbon enters the surface environment by weathering, volcanism and rock metamorphism and, over timescales longer than 100 Myr, $\delta_{in} = -5\%$, the average value for crustal carbon⁸. Where values of sedimentary δ_{carb} and δ_{org} can be measured, it is thus possible to determine f_{org} for ancient carbon cycles (all δ values in ‰)

$$f_{org} = (\delta_{in} - \delta_{carb})/(\delta_{org} - \delta_{carb}) = (\delta_{carb} + 5)/\Delta_c \quad (2)$$

The carbon isotopic record

Strauss *et al.*⁹ summarized 731 isotope analyses (Fig. 1) of total organic carbon in Proterozoic sediments. Except for some new low values for rocks older than 2.6 Gyr, the $\delta^{13}C$ values from earlier work and the new study do not differ substantially. Lines tracing the average values of δ_{carb} and δ_{org} in Fig. 1 are roughly parallel, creating the impression that Δ_c was invariant throughout the Proterozoic. Some authors have therefore emphasized aspects of uniformity in the Proterozoic carbon cycle and its isotope record^{10–12}. Episodic $\delta^{13}C$ variations have, however, been found^{13,14}.

The δ_{org} values in Fig. 1 reflect not only biological and palaeoenvironmental effects, but also postdepositional thermal degradation of the kerogen, which results in preferential loss of ^{13}C -depleted, hydrogen-rich products. As a result, changes in the H/C ratio and δ_{org} of residual kerogens are correlated¹⁵. As would be expected, older kerogens are more likely to be thermally altered and have lower H/C ratios¹⁶. Because δ_{org} has been affected, the true magnitudes of Δ_c and f_{org} during the Proterozoic have been obscured.

To minimize isotopic shifts due either to extreme thermal alteration or to analytical artefacts, Fig. 2 includes isotopic

compositions^{9,17} only for kerogens with ash contents less than 25% and H/C values greater than 0.1. Furthermore, the best view of the Proterozoic isotopic record can be obtained by reconstructing the values of δ_{org} before alteration. The characteristic trend relating changes in δ_{org} to values of H/C has been examined in detail^{15,18}. Relationships between changes in H/C and changes in δ_{org} are not significantly dependent on kerogen type or host lithology¹⁷. The curve which relates the most probable values for this shift ($\Delta\delta_{\text{org}}$) to H/C values (see Fig. 5-5 of ref. 17) has the form

$$\Delta\delta_{\text{org}} = 4.05 - 3.05r + 0.785/r + 0.0165/r^2 - (8.79 \times 10^{-4})/r^3 \quad (3)$$

where $\Delta\delta_{\text{org}} = \delta_{\text{org}}$ (as analysed, H/C = r) - δ_{org} (initial, H/C = 1.5). When initial δ values are reconstructed, the isotope record takes the form shown in Fig. 2. Because δ_{carb} is relatively constant, the changes in Δ_c must, as shown by equation (2), be linked principally to changes in f_{org} .

Burial of organic carbon

When the data of Fig. 2 are used to calculate values of f_{org} for 100-Myr increments between 2.5 and 0.6 Gyr, results indicate a long-term increase (Fig. 3). Complementary changes in δ_{carb} and δ_{org} (Fig. 2) support this trend. This increase was not monotonic but episodic; high values were attained between 2.1 and 1.8 Gyr and between 1.1 and 0.7 Gyr. These changes will have affected the total crustal inventory of organic carbon and, in turn, other aspects of the surface redox environment.

Computed values of f_{org} indicate the average degree of reduction of carbon being stored in accumulating sediments. At the extremes, $f_{\text{org}} = 0$ corresponds to no reduction (all carbon being buried as carbonate) and $f_{\text{org}} = 1$ corresponds to complete reduction (all carbon being buried as organic). This output of carbon

from the surface environment is balanced by inputs contributed by erosion and thermal activity. These carbon inputs will be at least partly reduced, because they include organic carbon being remobilized during destruction of older sediments. If the average degree of reduction of the output differs from that of the input, there must be a net transfer of oxidizing or reducing power from the carbon cycle to the cycle of some other element, presumably iron, sulphur or oxygen.

To reconstruct changes in the crustal inventory of organic carbon during the Proterozoic, we have begun by postulating that the f_{org} value observed at 2.6 Gyr accurately reflects carbon burial before that time. If so, the crust then contained about $0.09 \times 7.6 \times 10^{21}$ mole organic carbon ($= f_{\text{org}} \times C_{\text{tot}}$, where C_{tot} is the total number of moles of carbon in the crust¹¹). During the next 100 Myr, the time interval between 2.6 and 2.5 Gyr, a fraction of the crustal inventory of carbon will have been recycled, passing through the surface reaction chamber and being reburied with the f_{org} value characteristic of that epoch. Quantitatively, variations in the organic carbon reservoir can then be calculated as follows

$$M_{t-\tau} = M_t e^{-k\tau} + C_{\text{tot}}(1 - e^{-k\tau})f_{\text{org}} \quad (4)$$

where M is the quantity of organic carbon in the crust, t is the present age of sediments deposited at the beginning of an increment (t approaches 0 as time advances), τ is the duration of the increment, f_{org} is based on δ values observed during the interval t to $t - \tau$, and k is the first-order decay constant characteristic of the recycling of the sedimentary inventory. For a sediment half life of 400 Myr (ref. 19), $k = (\ln 2)/0.4 = 1.733 \text{ Gyr}^{-1}$. The first term on the right of equation (4) represents survival of crustal organic carbon from one time increment to the next, the second represents effects of recycling. Stepwise application of this equation yields the results summarized in Fig. 4.

Uncertainties in M can be discussed mainly in terms of the k parameter. First, k is certainly time-variant. Tectonic variations strongly influence rates at which crustal materials are recycled¹⁹. Rates of erosion and, thus, rates of burial of sedimen-

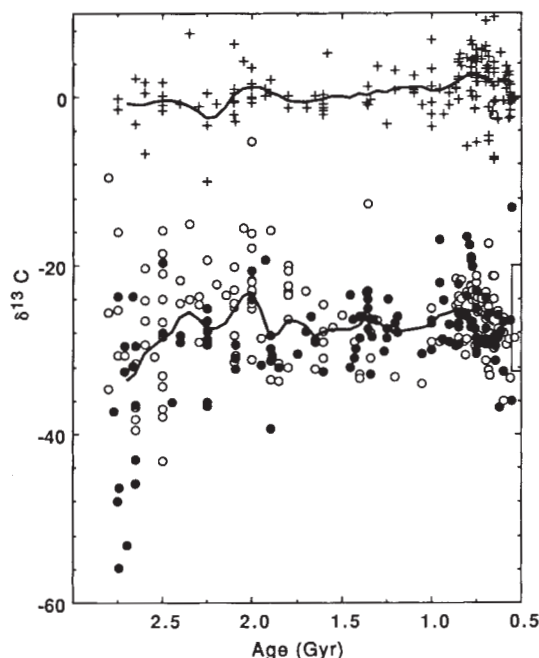


FIG. 1 δ_{carb} (crosses) and δ_{org} (circles) against age of unpurified kerogens. ●, Values obtained by Strauss *et al.*⁹; ○, earlier data (see Strauss *et al.*⁹ for compilation). Solid lines depict running averages, calculated for successive 100-Myr time increments, of 200 Myr intervals during the Proterozoic. For example, a δ value given for 1.0 Gyr is an average of values between 1.1 and 0.9 Gyr. The rectangle located midway along the right margin represents the range of δ_{org} values typically observed in Phanerozoic sediments⁶⁰.

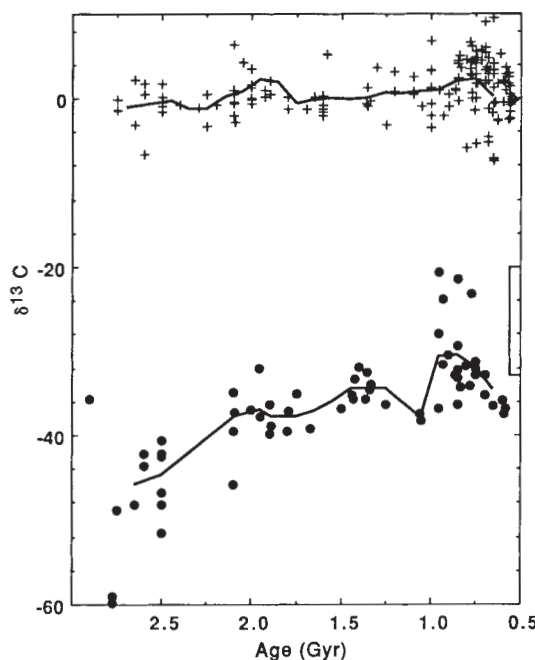


FIG. 2 δ_{carb} (crosses) and δ_{org} (circles) against age of purified kerogens; δ_{org} values are corrected for postdepositional alteration (see text). As in Fig. 1, the solid lines depict running averages of the data, and the rectangle along the right margin of the figure represents Phanerozoic δ_{org} values.

tary debris, are highest when high mountain terrain is extensive and/or when continental plates experience rifting. Rainfall and weathering are also important and are variable. For the late Proterozoic, the isotopic records of strontium and neodymium can be interpreted⁶ in terms of large variations in rates of erosion and burial. To examine the sensitivity of our estimates of M to variations in k , we have also considered sediment half lives of 0.3 Gyr and 0.5 Gyr (Fig. 4).

A second uncertainty related to k concerns the postulate that $M_{2.6}$ (the value 2.6 Gyr ago) can be estimated from f_{org} prevailing at that time. This estimate does seem at least to be representative of organic sedimentation during the interval 2.7 to 2.5 Gyr (Fig. 3). We have not used the entire isotopic record from before 2.5 Gyr to estimate variations in M from the beginning of the rock record, because the earlier record is incomplete and highly altered.

Values of δ_{carb} have been affected to some extent by postdepositional alteration of carbonates, yet no corrections have been made here. Because of their greater age, early Proterozoic carbonates might be more affected than late Proterozoic carbonates. But Veizer *et al.*²⁰ obtained a 'corrected' δ_{carb} value of $0 \pm 1.5\%$ for early Proterozoic marine carbonate, which is in reasonable agreement with the corresponding average δ_{carb} values in Fig. 2. Apparently, postdepositional alteration of carbonates has not seriously compromised the accuracy of the calculations of M .

Episodic release of oxidizing power

Assuming a half life of 0.4 Gyr we find that M increased in the Proterozoic mostly in two increments as follows: from 0.7×10^{21} to 1.05×10^{21} mol between 2.5 and 1.8 Gyr, and from 1.05×10^{21} to 1.3×10^{21} mol between 1.3 and 0.7 Gyr. It is instructive to explore both the causes of this trend in M and also its consequences for crustal and atmospheric abundances of oxidized species such as Fe^{3+} , SO_4^{2-} and O_2 . Rates of organic burial depend not only on the intensity of production of biomass but also on the preservation of organic matter as it is delivered to sediments and awaits burial. Productivity depends on the efficiency and environmental tolerance of photoautotrophs and on the delivery of nutrients to the photic zone^{3,21}. Organic

preservation depends on the versatility and aggressiveness of heterotrophs, on the availability of oxidants in the water column and sediments, and on the rate of sedimentation of inorganic debris, rapid burial assisting preservation^{22,23}. Thus both productivity and preservation are affected by both biological and geological processes. Although the evolutionary development of oxygenic photosynthesis must have enhanced productivity because it allowed autotrophs to make use of an abundant and ubiquitous electron donor, this evolutionary step occurred during the Archean²⁴⁻²⁷, at least 600 Myr before the time when, according to geochemical evidence, O_2 accumulated in the atmosphere. Thus the pulses of accelerated organic burial revealed by the Proterozoic record are expected to correlate with time intervals during which tectonically enhanced rates of weathering and erosion delivered more nutrients²⁸ and promoted rapid burial.

Most of the O_2 generated from the photosynthetic production of organic matter is consumed by respiratory processes. The small amount of organic matter that escapes oxidation by being buried (today, $\sim 0.2\text{--}0.3\%$ of primary production^{7,29}) allows an equivalent amount of O_2 to remain in the surface environment¹. Burial of organic carbon thus releases O_2 ; erosion commonly exposes organic carbon, sulphide, and Fe^{2+} which consume O_2 (refs 2, 3, 30). It is instructive to explore how this balance was affected by tectonic events during the Proterozoic.

Proterozoic events

The interval 3.0 to 2.4 Gyr probably saw the first assembly of large, relatively stable continental plates from smaller cratons³¹, setting the stage for extensive cratonic sedimentation³². Carbonate platforms having most of the essential features of their Phanerozoic equivalents were well developed by 2.6 to 2.3 Gyr (ref. 33). Indeed, the best-developed Proterozoic carbonate platforms, with evidence of extensive oxygenic photosynthetic stromatolitic communities, are primarily of early Proterozoic age³³. This evidence for extensive production of O_2 might seem to conflict with the observation from palaeosols that O_2 levels were low before 2.0 Gyr (ref. 34). But modern microbial mats remineralize organic matter efficiently (D. E. Canfield and D.J.D., manuscript submitted); the low amounts of organic carbon in early Proterozoic stromatolitic carbonates³⁵ are thus not unexpected. Oxidants can be consumed as rapidly as they are produced in these communities. Little organic carbon was buried on the extensive early Proterozoic carbonate platforms, so net accumulation of O_2 was near zero. This picture is consistent with the low values of f_{org} (0.1) observed for the earliest Proterozoic (Fig. 3). The smaller net production of oxidants was probably consumed by reactions in sea-floor hydrothermal systems³⁶, which were more active then than today. Accordingly, seawater SO_4^{2-} , although present, was substantially below modern concentrations³⁷. The deposition of Superior-type banded iron formations at that time apparently required an anoxic deep ocean³ containing some dissolved Fe^{2+} .

Commencing at 2.2–2.1 Gyr, the large continental plates that had assembled for the first time in the late Archean to early Proterozoic underwent rifting and, later, orogeny on a global scale, as evidenced by massive basic and ultrabasic dyke swarms and Andean-type orogenic belts³¹. Post-Archean tectonic cycles display an increasingly prominent early rift stage and a well-developed terminal stage of orogeny³². Accordingly, rifting on a global scale probably promoted the development of extensive anoxic basins favourable for organic preservation. A sustained rise in seawater $^{87}\text{Sr}/^{86}\text{Sr}$ values at that time³⁸ indicates enhanced rates of continental erosion. The earliest-known glaciations occurred during this interval³⁹, lowering the sea level, exposing more continent, and making conditions more favourable for high rates of erosion and clastic sedimentation⁴⁰. Between 2.1 and 1.7 Gyr, the opening and closing of Atlantic-type ocean basins caused rifting and uplift which accelerated the rates of erosion, release of nutrients for biological production,

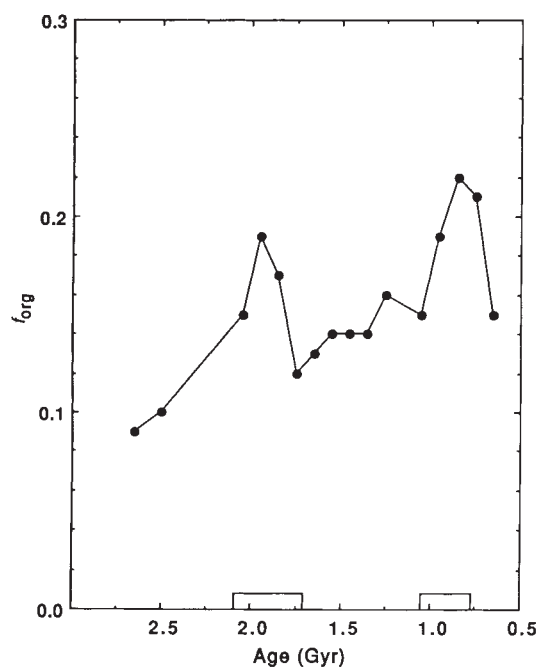


FIG. 3 Fraction of carbon buried as organic matter (f_{org}) against age of purified kerogens. Rectangles along the bottom margin depict time intervals of enhanced global rifting and orogeny (see text).

near-shore sedimentation and, therefore, organic burial. One such episode of enhanced organic sedimentation at this time might be represented by ^{13}C -rich carbonates reported by Baker and Fallick¹⁴.

The idea that growth of the crustal inventory of reduced carbon (M) led to oxidation of other elements is supported by geological evidence. A substantial rise in atmospheric O_2 levels is indicated by increase in retention of Fe^{3+} in palaeosols^{34,41} and emergence of extensive redbeds⁴². The biosynthesis of sterol precursors requires O_2 , and the oldest known sterane biomarkers occur in the 1.69 Gyr Barney Creek formation⁴³. The deposition of Superior-type banded iron formations was curtailed severely after 1.8 Gyr (ref. 44), perhaps because of oxygenation of the deep ocean. Possibly the oldest known occurrence of massive SO_4^{2-} evaporite deposition is recorded in 1.4–1.8-Gyr sediments of the McArthur basin, Australia⁴⁵.

During the interval 1.7 to 1.2 Gyr, orogenic activity was not as globally extensive as it was during the preceding period. There is no evidence of glaciation³⁹. A supercontinent may have existed, orogenic activity was more local in nature, and terms such as 'anorogenic magmatism' and 'abortive rifting' have been used to characterize major tectonic events of that time³¹. It is reasonable to conclude that global rates of sedimentation and organic burial were low to moderate. The absence of evidence for large changes in the sizes of the oxidized reservoirs is therefore not surprising. The O_2 content of the atmosphere was apparently maintained at a value intermediate between modern levels and those of the early Proterozoic⁴⁶.

Tectonic activity increased during the interval 1.2 to 0.9 Gyr. The emplacement of extensive mafic dyke swarms³¹ heralded the breakup of the supercontinent and the wide dispersal of continental fragments^{47,48}. Modern-style orogenic cycles³¹ and glacial episodes³⁹ became more frequent than before. Once again conditions were favourable for relatively rapid rates of organic sedimentation. Massive evaporites occur during and after this interval in central Africa (Upper Roan group⁴⁹),

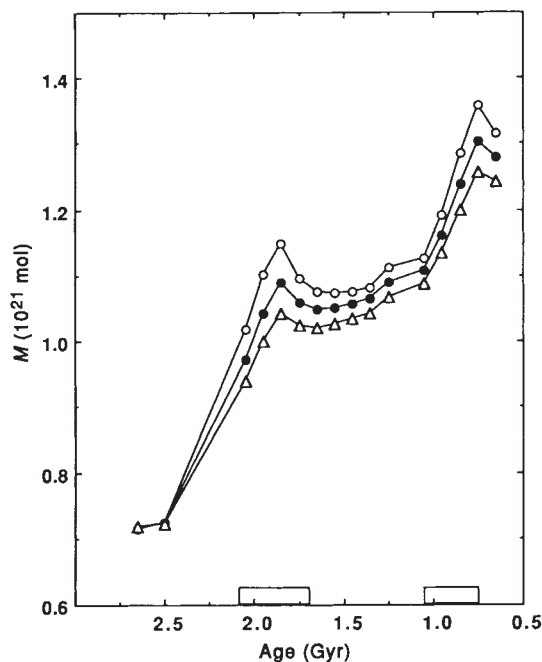


FIG. 4 Quantity of organic carbon in the crust (M) against age. Values of M are calculated according to equation (4) (in the text). Symbols represent calculations of M , assuming sediment half lives (see derivation of equation (4) as follows: $\circ$, 300 Myr; $\bullet$, 400 Myr; $\triangle$, 500 Myr. Rectangles along the bottom margin depict time intervals of enhanced global rifting and orogeny (see text).

North America (Grenville series⁵⁰) and Australia (Bitter Springs formation^{51,52}).

As indicated in Fig. 4, the crustal organic carbon reservoir increased in a stepwise fashion during each of the two time intervals during the early and middle-to-late Proterozoic, when global rifting and orogenies occurred. A third stepwise increase apparently occurred in a brief interval between the Varangian glacial episode and the Cambrian boundary⁶. These events are not correlated with the development of new biological sources of O_2 . Indeed, the last stepwise oxidation is correlated with the development of important new O_2 consumers, the metazoans. It is instead likely that these increases were driven by accelerated rates of erosion, nutrient release and clastic sedimentation, factors that are known to accelerate organic sedimentation on the modern Earth^{29,30}. These observations support earlier proposals that tectonic events strongly affected the biogeochemical carbon⁴⁰ and sulphur⁵³ cycles. That the oxidized crustal reservoirs of O_2 , SO_4^{2-} and Fe^{3+} also increased, as predicted², is supported by evidence which is particularly compelling for the interval 2.1 to 1.8 Gyr (ref. 41). A prediction from the present findings is that a second large increase in the oxidized reservoirs occurred between 1.1 and 0.8 Gyr. A systematic search for such evidence should be undertaken.

The role of life

These findings challenge the widely held view that innovations in biological evolution directed the long-term rise of atmospheric oxygen levels. As stated earlier, the development of oxygenic photosynthesis occurred at least 600 Myr before O_2 (according to geochemical evidence) accumulated in the atmosphere. Even eukaryotic organisms, which require O_2 for biosynthesis of essential lipids⁵⁴, appear in the palaeontological record about 2.1 Gyr ago⁵⁵, perhaps before the first large O_2 increase discussed here. Photosynthesis indeed provided an O_2 source strong enough to sustain a major atmospheric increase, but the timing and magnitude of O_2 accumulation was regulated by tectonic processes controlling erosion and sedimentation. These observations are consistent with the view^{19,56} that chemical properties of the environment that are buffered by large crustal reservoirs (such as crustal organic carbon and sulphur) over long time-scales (>10 Myr) will ultimately be controlled by geological processes, such as tectonics, which regulate interactions with those reservoirs.

The growth of the crustal organic carbon reservoir and the resultant oxidation of the surface environment must have profoundly affected the Earth's biota. Evidence of abundant eukaryotic organisms appears in the organic geochemical record about 1.7 Gyr ago⁴³. Because eukaryotes require O_2 , it is logical to associate an increase in their abundance and diversity with the rise of O_2 that probably accompanied the growth of the organic reservoir between 2.2 and 1.8 Gyr.

The declining $\delta^{13}\text{C}$ difference between sedimentary carbonate and reduced carbon during the Proterozoic (Fig. 2) suggests that isotopic discrimination during biological CO_2 uptake decreased in response to declining CO_2 concentrations in the oceans and atmosphere⁹. The fact that the greatest decline in discrimination coincides with the episodes of global rifting and orogeny is consistent with the drawdown of atmospheric CO_2 levels which accompanies high erosion rates⁵⁷. It is therefore likely that the atmospheric ratio of O_2 to CO_2 increased markedly during these episodes. Evolutionary changes in the enzyme ribulose biphosphate carboxylase oxygenase^{58,59}, which has both CO_2 and O_2 as substrates, might have been triggered by these transitions. $\square$

Received 12 May; accepted 14 September 1992.

- Holland, H. D. in *Hydrogeochemistry and Biogeochemistry* (ed. Ingerson, E.) 68–81 (The Clark, Washington DC, 1973).
- Garrels, R. M. & Perry, E. A. Jr in *The Sea* (ed. Goldberg, E.D.) 303–336 (Wiley, New York, 1974).
- Holland, H. D. *The Chemical Evolution of the Atmosphere and Oceans* 1–582 (Princeton Univ. Press, Princeton, 1984).

4. Degens, E. T., Guillard, R. R. L., Sackett, W. M. & Hellebust, J. A. *Deep-Sea Res.* **15**, 1–9 (1968).
5. Farquhar, G. D., O'Leary, M. H. & Berry, J. A. *Aust. J. Plant Physiol.* **9**, 121–137 (1982).
6. Derry, L. A., Kaufman, A. J. & Jacobsen, S. B. *Geochim. cosmochim. Acta* **56**, 1317–1329 (1992).
7. Betts, J. N. & Holland, H. D. *Palaeogeogr. Palaeoclim. Palaeoecol.* **97**, 5–18 (1991).
8. Holser, W. T., Schidlowski, M., Mackenzie, F. T. & Maynard, J. B. in *Chemical Cycles in the Evolution of the Earth* (eds Gregor, C. B., Garrels, R. M., Mackenzie, F. T. & Maynard, J. B.) 105–173 (Wiley, New York, 1988).
9. Strauss, H., Des Marais, D. J., Summons, R. E. & Hayes, J. M. in *The Proterozoic Biosphere: A Multidisciplinary Study* (eds Schopf, J. W. & Klein, C.) 117–128 (Cambridge Univ. Press, New York, 1992).
10. Schidlowski, M. in *Mineral Deposits and the Evolution of the Biosphere* (eds Holland, H. D. & Schidlowski, M.) 103–122 (Springer, Berlin, 1982).
11. Schidlowski, M. *Nature* **333**, 313–318 (1988).
12. Schidlowski, M., Hayes, J. M. & Kaplan, I. R. in *Earth's Earliest Biosphere* (ed. Schopf, J. W.) 149–186 (Princeton Univ. Press, Princeton, 1983).
13. Schidlowski, M., Eichmann, R. & Junge, C. E. *Geochim. cosmochim. Acta* **40**, 449–455 (1976).
14. Baker, A. J. & Fallick, A. E. *Nature* **337**, 352–354 (1989).
15. McKirdy, D. M. & Powell, T. G. *Geology* **2**, 591–595 (1974).
16. Strauss, H., Des Marais, D. J., Summons, R. E. & Hayes, J. M. in *The Proterozoic Biosphere: A Multidisciplinary Study* (eds Schopf, J. W. & Klein, C.) 95–100 (Cambridge Univ. Press, New York, 1992).
17. Hayes, J. M., Kaplan, I. R. & Wedeking, K. W. in *Earth's Earliest Biosphere* (ed. Schopf, J. W.) 93–134 (Princeton Univ. Press, Princeton, 1983).
18. Redding, C. E., Schoell, M., Monin, J. C. & Durand, B. in *Advances in Organic Geochemistry 1979* (eds Douglas, A. G. & Maxwell, J. R.) 711–723 (Pergamon, London, 1980).
19. Veizer, J. in *Chemical Cycles in the Evolution of the Earth* (eds Gregor, C. B., Garrels, R. M., Mackenzie, F. T. & Maynard, J. B.) 175–220 (Wiley, New York, 1988).
20. Veizer, J., Clayton, R. N. & Hinton, R. W. *Geochim. cosmochim. Acta* **56**, 875–885 (1992).
21. Broecker, W. S. *Scient. Am.* **249**, 146–161 (1983).
22. Muller, P. J. & Suess, E. *Deep-Sea Res.* **A26**, 1347–1362 (1979).
23. Canfield, D. E. in *NATO-ARW Interactions of C, N, P and S biogeochemical cycles* (eds Wollast, R., Chou, L. & Mackenzie, F.) (Springer, Berlin, in the press).
24. Buick, R. *Science* **255**, 74–77 (1992).
25. Beukes, N. J. & Lowe, D. R. *Sedimentology* **36**, 383–397 (1989).
26. Hayes, J. M. in *Earth's Earliest Biosphere* (ed. Schopf, J. W.) 291–301 (Princeton Univ. Press, Princeton, 1983).
27. Schopf, J. W., Hayes, J. M. & Walter, M. R. in *Earth's Earliest Biosphere* (ed. Schopf, J. W.) 361–384 (Princeton Univ. Press, Princeton, 1983).
28. Froelich, P. N., Bender, M. L., Luedke, N. A., Heath, G. R. & DeVries, T. *Am. J. Sci.* **282**, 474–511 (1982).
29. Holland, H. D. *The Chemistry of the Atmosphere and Oceans* 1–351 (Wiley, New York, 1978).
30. Berner, R. A. & Canfield, D. E. *Am. J. Sci.* **289**, 333–361 (1989).
31. Windley, B. F. *The Evolving Continents* 1–399 (Wiley, New York, 1984).
32. Lowe, D. R. in *The Proterozoic Biosphere: a Multidisciplinary Study* (eds Schopf, J. W. & Klein, C.) 67–76 (Cambridge Univ. Press, New York, 1992).
33. Grotzinger, J. P. in *Controls on Carbonate Platform and Basin Development* (eds Crevello, P. D., Wilson, J. L., Sarg, J. F. & Read, J. F.) 79–106 (Society of Economic Paleontologists and Mineralogists, Tulsa, Oklahoma 1989).
34. Holland, H. D., Feakes, C. R. & Zbinden, E. A. *Am. J. Sci.* **289**, 362–389 (1989).
35. Des Marais, D. J. *Palaeogeogr. Palaeoclim. Palaeoecol.* **97**, 93–96 (1991).
36. Veizer, J., Compston, W., Hoefs, J. & Nielson, H. *Naturwissenschaften* **69**, 173–180 (1982).
37. Eastoe, C. J., Gustin, M. S., Hurlbut, D. F. & Orr, R. L. *Precamb. Res.* **46**, 353–364 (1990).
38. Veizer, J. & Compston, W. *Geochim. cosmochim. Acta* **40**, 905–914 (1976).
39. Harland, W. B. *Mem. geol. Soc. Am.* **161**, 279–288 (1983).
40. Worsley, T. R., Nance, D. R. & Moody, J. B. *Paleoceanography* **1**, 233–263 (1986).
41. Holland, H. D. & Beukes, N. J. *Am. J. Sci.* **A290**, 1–34 (1990).
42. Eriksson, P. G. & Cheney, E. S. *Precamb. Res.* **54**, 257–269 (1992).
43. Summons, R. E. & Walter, M. R. *Am. J. Sci.* **A290**, 212–244 (1990).
44. Walker, J. C. G. et al. in *Earth's Earliest Biosphere* (ed. Schopf, J. W.) 260–290 (Princeton Univ. Press, Princeton, 1983).
45. Walker, R. N., Muir, M. D., Diver, W. L., Williams, N. & Wilkins, N. *Nature* **265**, 526–529 (1977).
46. Zbinden, E. A., Holland, A. D. & Feakes, C. R. *Precamb. Res.* **42**, 141–163 (1988).
47. Piper, J. D. A. *Earth planet. Sci. Lett.* **59**, 61–89 (1982).
48. Piper, J. D. A. *Mem. Geol. Soc. Am.* **161**, 11–34 (1983).
49. Fleischer, V. D., Garlick, W. G. & Haldane, R. (ed. Wolf, K. H.) 223–352 (Elsevier, Amsterdam, 1976).
50. Brown, J. S. & Engel, A. E. J. *Bull. geol. Soc. Am.* **67**, 1599–1622 (1965).
51. Lindsay, J. F. *Bull. geol. Soc. Am.* **99**, 852–865 (1987).
52. Stewart, A. J. *Sedimentology* **26**, 33–62 (1979).
53. Cameron, E. M. *Nature* **296**, 145–148 (1982).
54. Jahnke, L. & Klein, H. P. *J. Bacteriol.* **155**, 488–492 (1983).
55. Han, T. M. & Runnegar, B. *Science* **257**, 232–235 (1992).
56. Veizer, J. *Origins Life Evol. Biosphere* **18**, 13–39 (1988).
57. Berner, R. A., Lasaga, A. C. & Garrels, R. M. *Am. J. Sci.* **283**, 641–683 (1983).
58. Badger, M. R. & Andrews, T. J. in *Progress in Photosynthesis Research* (ed. Biggins, J.) III.9.601–III.9.609 (Martinus Nijhoff, Dordrecht, 1987).
59. Jordan, D. M. & Ogren, W. L. *Archives Biochem. Biophys.* **227**, 425–433 (1983).
60. Deines, P. in *Handbook of Environmental Isotope Geochemistry* (eds Fritz, P. & Fontes, J. C.) 329–406 (Elsevier, Amsterdam, 1980).

ACKNOWLEDGEMENTS. We thank D. Canfield, A. Fallick, J. Farmer, L. Jahnke & J. Veizer for reviews. This study was made possible by the Precambrian Paleobiology Research Group, coordinated by J. W. Schopf. The work was supported by grants from NASA's Exobiology Program.

A truncated activin receptor inhibits mesoderm induction and formation of axial structures in *Xenopus* embryos

Ali Hemmati-Brivanlou & Douglas A. Melton

Department of Biochemistry and Molecular Biology, Harvard University, 7 Divinity Avenue, Cambridge, Massachusetts 02138, USA

Activins can induce mesoderm in embryonic explants and have been proposed as the natural inducer in *Xenopus*. A mutant activin receptor that inhibits activin signalling is used to show that activin is required for the induction of mesoderm *in vivo* and the patterning of the embryonic body plan. Blocking the activin signal transduction pathway also reveals autonomous induction of a neural marker and unmasks a relationship between activin and fibroblast growth factor.

UNDERSTANDING the processes that lead from a fertilized egg to the formation of germ layers and subsequently to a body plan is a central goal of embryology. Much of what is known about the development of a vertebrate body plan comes from studies of amphibia where, at the tadpole stage, the main body axis consists of the dorsal structures notochord, spinal cord and somites organized anterior to posterior as head, trunk and tail. All animal tissues derive from the three germ layers and the mesoderm plays a pivotal role in organizing the body axis¹. Mesodermal cells lead the movements of gastrulation^{2,3}, are required for the patterning of the nervous system^{4,5}, and themselves give rise to the muscular, skeletal, circulatory and excretory systems. Moreover, a portion of the dorsal mesoderm from early gastrula, the Spemann organizer, can induce and organize a second body axis following transplantation to another site⁶. An understanding of the development of mesoderm will help to clarify how the vertebrate body plan is generated.

Before gastrulation the three germ layers are simply arranged, top to bottom, in a frog blastula. Ectoderm arises from the top, or animal pole; mesoderm from the middle, or marginal zone, and endoderm from the bottom or vegetal pole. Mesoderm can be induced in animal pole cells (animal caps) by signals emanating from the vegetal pole⁷. Several peptide growth factors have been identified that can induce mesoderm in animal caps *in vitro*. When animal cap tissue is explanted from a blastula embryo and cultured in isolation it develops into a ball of epidermis. But in the presence of an inducing factor, the animal cap will differentiate into mesodermal derivatives, including notochord, muscle and blood^{8,9}. Members of the fibroblast growth factor family, in particular basic fibroblast growth factor (bFGF)^{10,11}, and the transforming growth factor- β (TGF- β) family¹², notably activins^{13–15}, are potent inducers in this assay. *Xenopus* homologues of the *Wnt* gene family may also have a role in mesoderm induction. Both *Xwn1* (ref. 16) and *Xwn18*

messenger RNAs elicit dorsal mesoderm formation when injected into the ventral side of an early embryo¹⁷⁻¹⁹, an activity shared to a lesser extent by activin RNA²⁰. bFGF and activin protein²¹ can be detected in the early embryo and although there are no data on the localization of activin, there is evidence that bFGF is present in the marginal zone and vegetal pole of early blastula²². Although Xwnt1 and Xwnt8 are not present at the proper time or place to effect dorsal mesoderm induction, there may be other Xwnts that fulfil this role.

Taken together, the experiments described above indicate that FGFs, activins and Wnts may have a role in mesoderm induction *in vivo*, but their exact function and relative importance is quite uncertain. For example, it is not clear how different mesodermal tissues and their spatial patterns could be specified by these factors, although there is evidence that the type^{23,24} and concentration²⁵ of the inducing molecule could be important. At the same time, studies have shown that the cells responding to the inducer have an inherent bias or prepattern that can determine which type of mesoderm will be formed^{26,27}. Perhaps the most pressing question is whether any of these factors functions *in vivo* to induce mesoderm or merely mimic the natural inducer(s).

As specific genes cannot be removed by conventional genetic means in amphibians, alternative approaches must be tried²⁸. For example, a dominant negative mutant has been used to demonstrate a role for FGF in mesodermal development²⁹. Expression of a dominant negative FGF receptor perturbs early development, particularly the development of trunk and tail mesoderm. We have followed this approach to investigate the *in vivo* role of activins in mesoderm induction. Two maternally expressed *Xenopus* activin receptors (XAR1 and XAR7)³⁰⁻³² have been identified, each of which contains extracellular, single transmembrane and cytoplasmic serine/threonine kinase domains. These receptors bind activin with equal affinity, but not TGF- β 1 (ref. 31; and C. Kinter, personal communication). We have constructed a truncated version of XAR1 that contains the entire extracellular and transmembrane domains but which lacks nearly all of the cytoplasmic domain, including the serine/threonine kinase (Fig. 1a). By analogy with signal transduction by tyrosine kinase receptors, like the FGF receptor, we postulate that a truncated activin receptor (tAR) will interfere with the function of endogenous activin receptors by creating an inactive receptor heterodimer. Alternatively, if dimerization

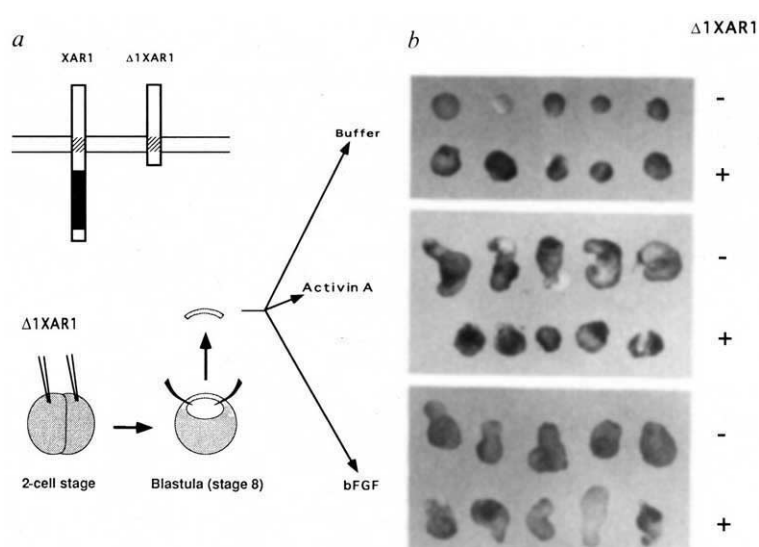
is not necessary for receptor function, the mutant receptor might compete with the endogenous receptor for activin binding.

Mesoderm induction in explants blocked by tAR

Synthetic RNA encoding the truncated activin receptor (Δ 1XAR1; Fig. 1) was injected into the animal pole at the two-cell stage and the effects on isolated animal caps were assessed by examination of gross morphology (Fig. 1b) and the presence or absence of seven tissue-specific molecular markers (Fig. 2). Late blastula animal caps were explanted and treated with bFGF or activin as inducing agents. Previous studies have shown that activin, and to a lesser extent bFGF, induces mesoderm in animal caps, an early sign of which is the morphogenetic elongation of animal caps³³. In the case of activin, these movements will ultimately result in formation of 'embryoids' with a full spectrum of differentiated mesodermal tissues and antero-posterior axial patterning¹⁵. Figure 1b shows that the truncated receptor completely and specifically blocks the early morphogenetic response of animal caps exposed to activin. Notably, the truncated receptor does not affect the response to bFGF. An unexpected observation is that animal caps derived from Δ 1XAR1 injected embryos, including those incubated in buffer alone, form cement glands. Control and uninjected animal caps do not form cement gland or mesodermal tissues.

Early and late molecular markers for mesoderm induction are also specifically blocked by the truncated activin receptor. In animal caps treated with activin, *Xenopus* brachyury (Xbra)³⁴ and Xhox-4, a homeobox protein closely related to Mix-1 (ref. 35), serve as immediate early markers for mesoderm induction. When animal caps injected with the Δ 1XAR1 RNA are incubated with activin, the expression of both Xbra and Xhox-4 is blocked (Fig. 2a). The specificity of this block is demonstrated by the fact that induction of Xbra by bFGF is not inhibited. In fact, Xbra induction by bFGF is enhanced in the presence of Δ 1XAR1 (Fig. 2a). In midgastrulae, brachyury is normally present as a ring in the lower part of the marginal zone (ref. 34, and Fig. 3b) and is a marker for prospective dorsal, lateral, as well as ventral mesoderm. Xhox-4, like Mix-1, marks the early vegetal cells as well as the prospective dorsal and ventral mesoderm. Therefore, the inhibition of Xbra and Xhox-4 expression suggests that all types of mesoderm are blocked by the truncated receptor. Further support for this claim comes from the fact that gooseoid³⁶, a marker for head mesoderm at

FIG. 1 The truncated activin receptor blocks the morphogenetic movements of animal caps exposed to activin, but not FGF. **a**, Wild-type (XAR1) and truncated activin receptor construct (Δ 1XAR1). Hatched boxes represent the transmembrane domains and the black box represents the cytoplasmic serine/threonine kinase domain. **b**, Assay of Δ 1XAR1 function in animal caps. Fertilized eggs were injected at the two-cell stage into the animal pole of each blastomere and animal caps were dissected from blastulae for culturing with or without inducing factors. Δ 1XAR1-injected animal caps (+) and globin-injected controls (−) were incubated with 50 pM activin, 3 nM bFGF, or buffer alone. The elongation movements that accompany induction by peptide growth factors mimic the convergent extension of mesoderm during normal development³³. **METHODS.** To construct Δ 1XAR1, a fragment of DNA from XAR1 (ref. 32) encoding the entire extracytoplasmic domain (including the signal sequence), the transmembrane domain and 10 amino acids from the cytoplasmic domain (excluding the serine/threonine kinase domain) and entirely free of 5' and 3' untranslated sequences was subcloned into pSP64T (ref. 48). The linearized plasmid was transcribed *in vitro* with SP6 to generate capped sense RNA. Eggs were obtained from female *Xenopus laevis* and embryos were raised as described⁴⁹. Either Δ 1XAR1 or *Xenopus* β -globin mRNA (2 ng) was injected into the animal pole of each blastomere of two-cell embryos that were incubated in 3% Ficoll, 0.5 $\times$ MMR buffer (1 $\times$ MMR is 5 mM HEPES, 100 mM NaCl, 2 mM KCl, 1 mM MgSO₄, 2 mM CaCl₂ and 1 mM EDTA, pH 7.8). Injected animal caps were isolated at stage 8–9 (ref. 50) frog blastula and incubated in either buffer alone



(0.5 $\times$ MMR), 50 pM purified activin A (by courtesy of G. Mathers, A. Mason and R. Schwall), or 3 nM *Xenopus* recombinant bFGF (gift of D. Kimelman). After 8–10 h, caps were examined for elongation by microscopy. Results shown are invariable for $n=70$ caps per condition.

the midgastrula stage, was absent in all injected caps (data not shown), and the induction of *Xwnt-8* expression, a marker for ventral mesoderm^{17,19}, in response to activin, is blocked by injection of the truncated activin receptor (Fig 2a, bottom panel).

Expression of muscle actin, a mesoderm-specific gene that is expressed at the end of gastrulation and increases during neurulation³⁷, is also selectively inhibited in animal caps injected with the truncated receptor (Fig. 2b). The block to induction of muscle actin is dependent on the dose of $\Delta 1XAR1$ (Fig. 2d). It is interesting to note that bFGF induces muscle actin to roughly 10-fold higher levels in caps injected with the truncated activin receptor compared with control caps (Fig. 2b, compare lanes 3 and 6). These data, and the enhancement of *Xbra* expression

in response to bFGF in the presence of $\Delta 1XAR1$ (Fig. 2a), can be interpreted to indicate that activin antagonizes the mesoderm-inducing capacity of bFGF and that the mutant receptor, by blocking the effect of activin, amplifies or unmasks additional inducing activities of bFGF. This observation demonstrates a functional redundancy in embryonic cells whereby a block in one signalling pathway can lead to the enhancement of a parallel pathway to compensate for the effect.

The results presented in Fig. 2 show that the truncated activin receptor specifically blocks the function of the endogenous receptor and thus behaves as a dominant negative mutant²⁸. A stringent test for the specificity of a dominant negative mutant is the ability to rescue the phenotype by providing more

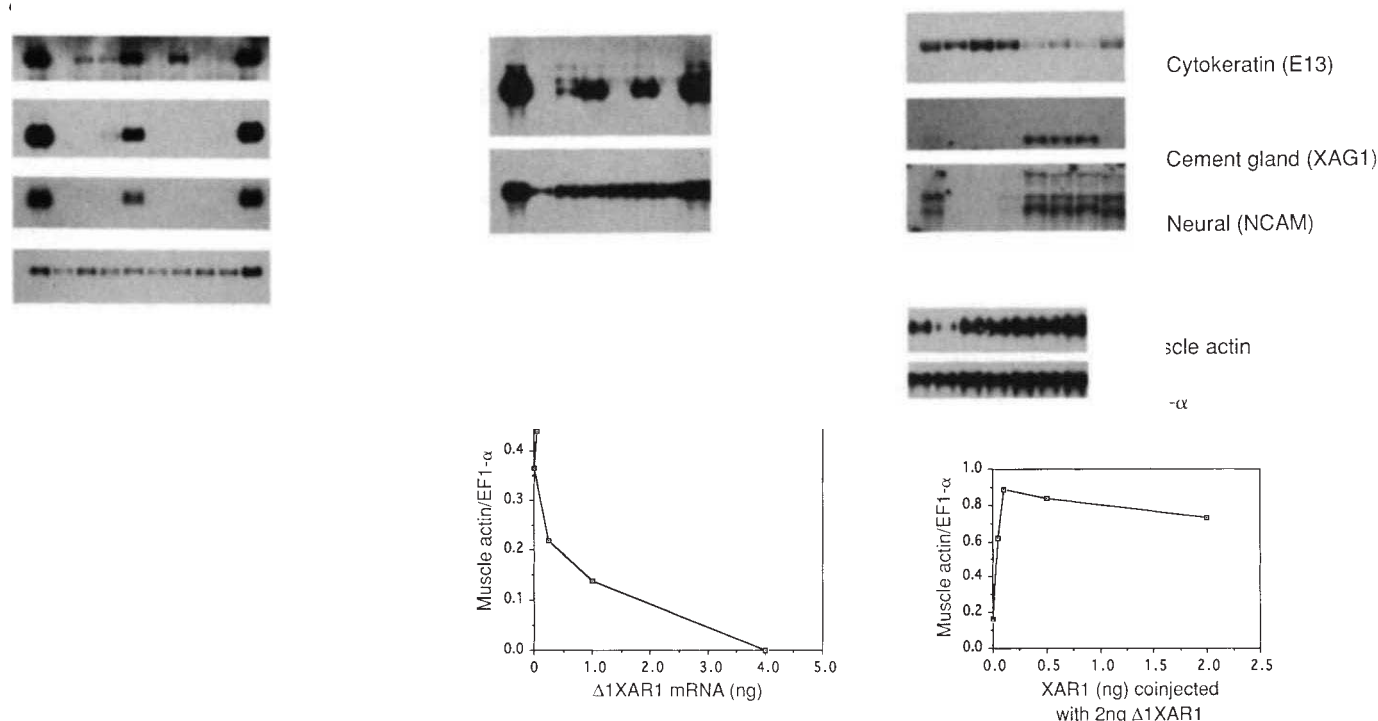


FIG. 2 The truncated activin receptor blocks immediate early and late mesodermal markers and induces neural-specific gene expression in animal caps. *a*, Immediate early markers. Embryos were injected with globin (control) RNA (lanes 2–5) or $\Delta 1XAR1$ RNA (lanes 6–9). Animal caps were dissected at the blastula stage and incubated either with buffer alone (lanes 2 and 6), bFGF (lanes 3 and 7), activin (lanes 4 and 8) or P388D1-inducing factor (PIF), a conditioned medium containing activin¹⁵ (lanes 5 and 9) for 3 h (midgastrula equivalent). Total RNA from 10 animal caps (lanes 2–9) was isolated and assayed for the presence of brachyury, *Xhox-4* and *Xwnt8* transcripts by northern blot analysis. Lane 1 contains total RNA from whole uninjected embryos at midgastrula. Lane 10 contains RNA from whole embryos injected in the animal pole with $\Delta 1XAR1$ RNA. This control shows that the truncated receptor injected in the animal pole does not effect gene expression in the marginal zone or vegetal pole of the whole embryo. The same northern blot was subsequently stripped from its signal and rehybridized with a probe for EF1- α to show that a comparable amount of RNA was loaded in each lane. *b*, Expression of a late mesodermal marker, muscle actin, is specifically blocked by $\Delta 1XAR1$. Animal caps were removed at the blastula stage from control (lanes 2–4) or $\Delta 1XAR1$ RNA (lanes 5–7) injected embryos as above and were incubated with either buffer alone (lanes 2 and 5), bFGF (lanes 3 and 6), or activin (lanes 4 and 7). Lane 1 is total RNA from tail-bud-stage control embryos. Lane 8 shows embryos at the same stage injected with $\Delta 1XAR1$ in the animal pole but from which the animal cap was not removed; this shows that $\Delta 1XAR1$ injected in the animal pole does not effect the expression of actin expression in the marginal zone. *c*, Effect of the truncated receptor on ectodermal derivatives. The same filter used in *b* was stripped and rehybridized successively with probes for either an epidermal marker, cytokeratin (E13)^{38,39}, a cement gland-specific marker (XAG1)³⁸ or a neural-specific gene, NCAM⁴⁰. *d*, The block of mesoderm

induction in animal caps injected with $\Delta 1XAR1$ is dose-dependent. Animal caps injected with different doses of $\Delta 1XAR1$ are incubated in the presence of the potent mesoderm inducer PIF for 12 h and analysed for muscle actin RNA. In each sample the amount of the translational elongation factor EF1- α is used as an internal control and the ratio of muscle actin signal to EF1- α is plotted versus the amount of $\Delta 1XAR1$ RNA injected. Complete blockage of muscle actin expression is obtained with 4 ng $\Delta 1XAR1$ RNA. *e*, Top, rescue of the competence of animal caps to respond to mesoderm induction by activin. Two-cell embryos were coinjected with 2 ng $\Delta 1XAR1$ and from 0 to 2 ng wild-type XAR1 activin receptor. Animal caps removed from blastula were incubated with activin and assayed for the amount of muscle actin and EF1- α by northern blot analysis (top panels). Animal caps from globin injected embryos are shown in lane 1. Animal cap RNA from embryos coinjected with 2 ng $\Delta 1XAR1$ and 0 ng (lane 2), 0.05 ng (lane 3), 0.1 ng (lane 4), 0.5 ng (lane 5) and 2 ng XAR1 (lane 6). Lane 7 is animal cap RNA derived from embryos injected with 2 ng XAR1 alone in the animal pole. Bottom, quantitation of data shown above. The addition of 50 pg wild-type XAR1 is sufficient to restore the full competence of animal caps to respond to activin. **METHODS.** Total RNA preparation and northern blotting have been described⁵¹. Probes were purified DNA fragments labelled by random priming and were used at 5×10^7 c.p.m. per blot. The brachyury (*Xbra*) probe was a gift from P. Wilson. The *Xhox-4* gene is an immediate early marker closely related to *Mix-1*³⁵ (P. Vize, personal communication). EF1- α was described by Krieg *et al.*⁵². E13 is a *Xenopus laevis* cDNA clone that cross-hybridizes to the previously described cytokeratin RNA, DG81^{39,53}. XAG1 is a cement gland-specific probe³⁸, and NCAM is a neural-specific gene⁴⁰. All control caps, unless otherwise mentioned, are injected with 2 ng globin RNA per blastomere. Intensity of radioactive bands on northern blots was quantified using a Fuji phosphorimager.

wild-type gene product as shown in Fig. 2e. When 50 pg XAR1 is coinjected with 2 ng Δ 1XAR1, it restores the full responsiveness of animal caps to activin.

Expression of ectoderm-specific genes

The effects of expression of the truncated activin receptor on ectodermal differentiation was examined using three molecular markers. After 12 h of incubation, with or without inducing factors, cement gland is detected in animal caps injected with Δ 1XAR1, as shown by the expression of a cement gland marker (XAG1 (ref. 38); Fig. 2c). An epidermal cytokeratin transcript (E13)^{38,39} is expressed in all samples, although its level is somewhat reduced in animal caps expressing the truncated receptor. Most surprising, however, is the finding that the neural cell adhesion molecule NCAM, a neural-specific marker⁴⁰, is expressed at high levels in animal caps injected with the truncated activin receptor. Indeed, there is more NCAM in the Δ 1XAR1-injected caps incubated in buffer alone than in the control caps induced with activin, which form abundant neural tissue^{15,20,41} (Fig. 2c). Several lines of evidence had demonstrated that activin is not a direct neural inducer. For example, activin does not induce dispersed animal cap cells to become neural, but instead dispersed cells become neural inducers after activin treatment²⁴. Moreover, dorsal mesoderm can induce neural tissue in animal cap explants at a later stage than can activin. It was inferred that activin induces dorsal mesoderm, which in turn induces neural tissue^{25,42,43}. Our data strongly suggest that neural-specific gene expression can be uncoupled from the formation of mesoderm in intact animal caps. Moreover, these results raise the intriguing possibility that activin signalling must be inhibited for neural induction to occur. A more detailed analysis of neural induction by the truncated activin receptor will be reported elsewhere (A.H.-B. and D.A.M., manuscript in preparation).

Endogenous mesoderm inducing signal(s) inhibited

Our experiments have shown that a truncated activin receptor can block the induction of mesoderm in explanted animal cap cells. In whole embryos, mesoderm is derived from the marginal zone (the equatorial region of the blastula) rather than from animal cap cells, which normally follow an ectodermal fate. To test whether the truncated receptor can block induction of mesodermal markers in the marginal zone of intact embryos, the truncated receptor was injected into one cell of two-cell embryos. In each case the uninjected half of the embryo serves as a control. Figure 3 shows that the expression of brachyury, which is expressed as a complete ring in control embryos, is reduced to a half ring in injected embryos. This indicates that the truncated activin receptor blocks the induction of Xbra RNA in cells that form mesoderm *in vivo*.

Δ 1XAR-expressing phenotype of whole embryos

The experiments described suggest that it should be possible to determine the phenotype of whole embryos lacking mesoderm and its accompanying morphogenetic movements. Embryos injected with the truncated activin receptor in both cells at the two-cell stage display a range of phenotypes, all of which are markedly deficient in mesodermal and axial development (Table 1; Fig. 4). By the time controls reach the tail-bud stage, injected embryos showing the extreme phenotype (about 50% of those injected; see Table 1) are grossly deficient in their development. In the most extreme cases there is no sign of body axis formation and although the embryos appear to have retained an animal-vegetal axis, there is little indication of an anterior-posterior or dorsal-ventral body plan (Fig. 4a). Occasionally some bottle cells form and gastrulation begins, but invagination rarely proceeds beyond a small lip (Fig. 4c). As observed in injected animal cap explants (Fig. 2), cement gland differentiation does occur (Fig. 4b). Histological sections reveal no evidence of mesodermal differentiation or gastrulation in the extreme cases

TABLE 1 Axial defect produced in embryos injected with Δ 1XAR1

RNA injected	Δ 1XAR1		β -Globin	
Phenotype	n	Per cent	n	Per cent
No axial structures	40	53	0	0
Partial axial defects	18	24	2	2
Normal embryos	17	23	92	98
Number of embryos scored	75	100	94	100

Embryos were injected with 4 ng of either Δ 1XAR1 or β -globin RNA in the equatorial region of both blastomeres of the two-cell embryo. Embryos were allowed to develop until sibling uninjected controls reached the tail-bud stage, at which time the survivors were scored for their phenotypes. Of the embryos without axial structures, roughly half showed minimal evidence of gastrulation (scant bottle-cell formation only) and had no detectable muscle or notochord either by histology or by molecular marker assay. The others showed only partial gastrulation and contained less than 20% of the normal amount of these markers (Fig. 4e). Partial axial defects are animals without heads or tail and with less than 50% of the normal amount of notochord and muscle. This division into classes, extreme and partial defects is arbitrary in that a range of defect is observed. n, Number of embryos.

(Fig. 4b). The blastocoel remains intact and there is virtually no rearrangement of the presumptive germ layers. Assays for molecular markers of differentiation show that these embryos do not form notochord, express muscle actin or Xbra RNA (Fig. 4f). The absence of mesodermal differentiation, as assayed by muscle actin and Xbra RNA expression, is dependent on the dose of truncated activin receptor RNA injected. At 4 ng of injected RNA, muscle actin and brachyury expression are both drastically reduced (Fig. 4f).

The rest of the embryos show, to varying degrees, rudimentary signs of mesodermal development, gastrulation and axis formation, but normal embryos are not produced (Fig. 4d, and Table 1). There is always a marked reduction in notochord and muscle formation, with most of the embryos in this class forming less than half the normal amount of notochord or muscle (Fig. 4e).

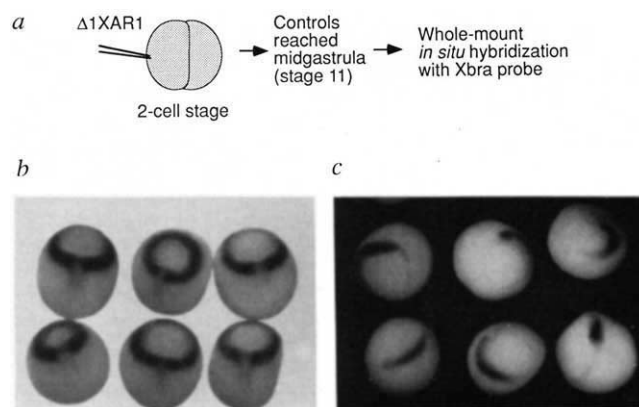


FIG. 3 The truncated activin receptor blocks the induction of the mesodermal marker brachyury in the marginal zone cells *in vivo*. a, Experimental design: Δ 1XAR1 or control globin RNA were injected in the equatorial region of one cell at the two-cell stage. Embryos were allowed to grow until sibling controls reached the midgastrula stage (stage 11) and assayed for brachyury expression by whole-mount *in situ* hybridization. b, The torus of brachyury expression in globin (control) injected embryos is undistinguishable from uninjected embryos. c, Embryos injected with Δ 1XAR1 block the induction of brachyury expression in only one side. Whole-mount *in situ* was performed on more than 25 injected embryos and these were all found to express brachyury only on their uninjected side. METHODS. Albino embryos were injected with 2 ng of either Δ 1XAR1 or globin RNA. Whole-mount *in situ* hybridization was done as described^{54,55} using an anti-sense brachyury probe (gift of P. Wilson) synthesized in the presence of digoxigenin-UTP.

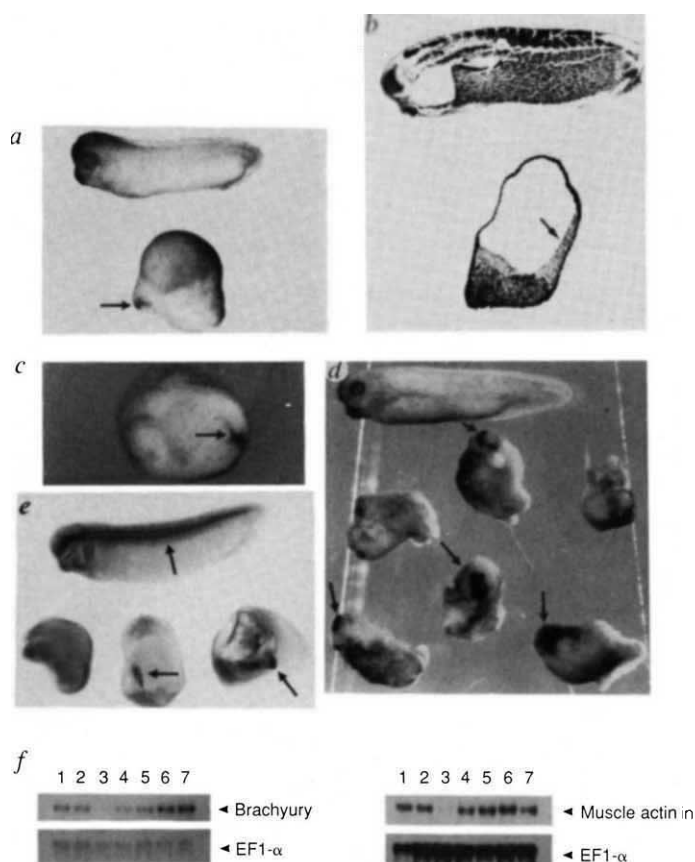


FIG. 4 Phenotype of embryos injected with $\Delta 1XAR1$. *a*, Phenotype of control and $\Delta 1XAR1$ -injected embryos. Top, a tailbud embryo (stage 27) injected with control RNA at the two-cell stage; bottom, equivalent-stage embryo injected with $\Delta 1XAR1$ RNA and representative of the extreme phenotype (see also Table 1). No axial structures or differentiated tissues are present, with the exception of cement gland (arrow). *b*, Longitudinal section through a control (top) and a $\Delta 1XAR1$ -injected embryo (bottom) at equivalent stages. Axial and mesodermal tissues present in the control embryo are missing in the injected embryo. The blastocoel cavity has expanded such that the roof of the cavity is a single cell layer thick. A few cells have migrated on the wall of the blastocoel (arrow), but these too have failed to differentiate. *c*, Vegetal view of $\Delta 1XAR1$ -injected embryo showing that some bottle cells have formed at one side (arrow), but invagination has not proceeded to form a full blastopore ring. *d*, Less extreme phenotypes at the equivalent of tail bud, stage 27 (see also Table 1), showing partial formation of aberrant axes and cement gland differentiation (arrows), although neither an anterior-posterior nor a dorsal-ventral axis can be readily recognized. *e*, Whole-mount immunohistochemistry with a notochord antibody. Top, control embryo showing the normal amount of notochord (arrow) present at the tail-bud stage. Bottom, three examples of embryos showing the extreme phenotype (see also Table 1), including one in which notochord is absent (left) and two with very little (<20%; see arrows) and aberrant notochord (right). *f*, Dose response effect of brachyury and muscle actin expression in embryos injected with $\Delta 1XAR1$ RNA. Embryos injected with different doses of the truncated receptor were allowed to develop until sibling controls reached mid-gastrula stage for brachyury expression, or tail-bud stage for muscle actin expression by northern blot. Lane 1, uninjected embryo; lane 2, embryos injected with 4 ng control globin RNA; lanes 3–7, embryos injected with 4 ng, 1 ng, 0.25 ng, 0.06 ng or 0.015 ng $\Delta 1XAR1$ RNA.

METHODS. Embryo embedding and histology were performed as previously described²⁰. The notochord-specific monoclonal antibody, Tor 70.1 (ref. 39), a gift from P. Kushner, was used at 1:200 and whole-mount immunohistochemistry was done as described⁵⁶. *Xenopus* globin or $\Delta 1XAR1$ RNA (10 nl of a 0.2 mg ml⁻¹ solution) were injected equatorially into both blastomeres of two-cell stage embryos. For each dose of $\Delta 1XAR1$, five embryos displaying the most extreme phenotypes were pooled for RNA preparations, northern blots and quantitation (as described in Fig. 2).

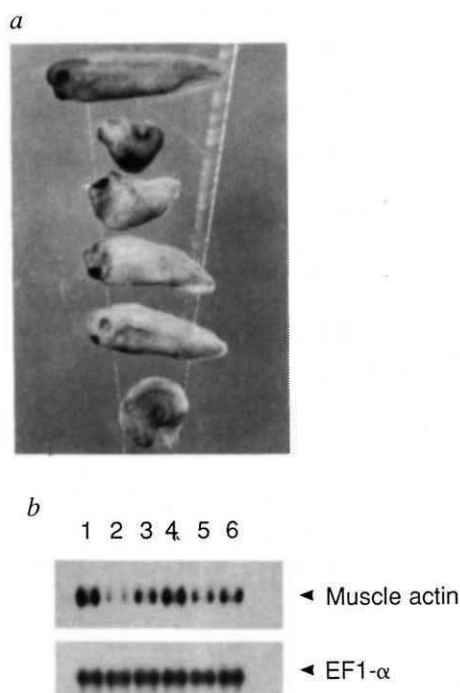


FIG. 5 Rescue of the $\Delta 1XAR1$ phenotype by wild-type XAR1 RNA. *a*, Dose-response effects of $\Delta 1XAR1$ and XAR1 RNA coinjection. The top embryo is injected with 4 ng globin RNA and shows no phenotypic abnormality. The remaining embryos, from top to bottom, were coinjected with the following ratios (in ng per blastomere) of $\Delta 1XAR1$ to XAR1: 2:0; 2:0.05; 2:0.25; 2:1 and 2:2. Note that when 0.25 or 1 ng XAR1 is coinjected with 2 ng $\Delta 1XAR1$ RNA, the embryos display a normal phenotype. At higher concentrations of XAR1 the embryos display an XAR1-specific phenotype characterized by multiple axes and multiple axial pattern^{30–32}. *b*, Dose-response effect of XAR1 rescue on the expression of muscle actin. The same set of injections described in *a* were given and embryos were collected at tail-bud stage and assayed for muscle actin RNA by northern blot analysis. Expression of the mesodermal marker is rescued by 0.05 ng XAR1 coinjected with 2 ng $\Delta 1XAR1$.

METHODS. Embryos were injected equatorially in both blastomeres of the two-cell stage. RNA from extreme phenotypes were isolated and analysed on northern blots as already described.

None of these mesodermal or axial defects are observed to any significant degree in embryos injected with control RNAs (Fig. 4f, and Table 1).

There are at least two explanations for the different classes of embryos formed in the presence of the truncated activin receptor. We believe the extreme phenotype, embryos showing

no mesoderm or anterior-posterior/dorsal-ventral axes, is due to the effective elimination of activin signalling by ectopic expression of the truncated receptor. The second class of embryos, those with some mesodermal tissues, could arise because the injected RNA did not fully diffuse throughout the marginal zone before cleavage divisions took place⁴⁴.

Alternatively, embryos may be able to compensate or regulate for the absence of mesodermal induction by activin, perhaps by using a second pathway. Consistent with this latter interpretation is the finding that FGF activity is enhanced in animal caps in the presence of the truncated activin receptor (Fig. 2). But even if this compensation does ensue, it is evidently not sufficient to allow for a full rescue of mesodermal development.

Rescue by wild-type activin receptor

If disruption of activin signalling by the truncated activin receptor is the cause of the mesodermal and axial deficiencies observed, then injection of wild-type activin receptor should rescue the phenotype. This has been tested by injecting increasing amounts of wild-type activin receptor RNA with a constant amount of RNA encoding the truncated receptor. The wild-type receptor can rescue embryos, as judged by gross morphology and molecular assays for mesodermal markers (Fig. 5). It is interesting to note that the rescue requires a relatively low amount of wild-type activin receptor and that higher concentrations of the wild-type receptor generate multiple and bent axes, as is observed by ectopic expression of the wild-type receptor alone^{30–32}.

Discussion

The experiments reported here show that disruption of activin signalling can, in the extreme phenotype, prevent mesoderm induction and dorsal body axis formation. Histological differentiation of mesodermal tissue is missing, early and late molecular markers are blocked and gastrulation does not proceed. These results establish activin as a principal determinant of mesoderm induction *in vivo* and not merely a molecule that mimics some other inducer.

The truncated activin receptor does not block mesoderm induction by exogenous FGF in animal caps, and yet endogenous FGF does not induce mesoderm in a significant fraction of embryos injected with $\Delta 1XAR1$. Furthermore, when half the embryo is injected with $\Delta 1XAR1$, that half lacks Xbra expression in all embryos tested ($n = 25$), even though FGF is

a potent inducer of brachyury in the animal cap assay³⁴. These data show that endogenous FGF signalling is not sufficient to rescue brachyury expression or mesoderm induction in the marginal zone of embryos injected with $\Delta 1XAR1$. At the same time, it is clear from other experiments with a dominant negative FGF receptor that FGF plays an important role in axial patterning, particularly for posterior structures²⁹. Taken together, these findings raise the interesting possibility that FGF signalling at the time of mesoderm induction requires a functional activin pathway. A confounding observation is that many embryos injected with $\Delta 1XAR1$ do form some mesoderm, albeit at reduced levels. Whether this is due to incomplete penetrance of $\Delta 1XAR1$ or endogenous FGF or of other activities, for example Wnt, is not yet clear.

An unexpected finding reported here is that the block in activin signal transduction autoinduces cells of the animal cap to switch to a neuronal fate in the absence of any detectable mesoderm. Presumptive epidermal in animal caps can respond to activin and form mesodermal tissues, but if this mesodermal specification by activin is blocked, the cells become neural. We therefore propose that an endogenous activity that blocks activin signalling is a candidate for a neural inducer. Two such molecules have been characterized that might fulfil this function. Follistatin interferes with activin signalling by binding the ligand⁴⁵, and inhibin can compete with activin for binding to the activin receptor³¹.

Although our study provides direct evidence for an inductive function of activin in frog embryogenesis, important questions are still unanswered. It remains to be determined precisely when and where activin signalling occurs *in vivo*. Also the isolation and characterization of other members of TGF- β family of receptors, such as the ones for the decapentaplegic-Veg1-related (DVR) protein family^{46,47}, will allow us to determine whether the truncated activin receptor interferes with the signalling of related receptors in these assays. Similarly, the relative roles of activins, Wnts and FGF is still unclear. Studies in other vertebrates should show whether or not the results obtained with frogs are generally applicable. □

Received 12 June; accepted 10 September 1992.

- Keller, R. in *Methods in Cell Biology* (eds Kay, B. K. & Peng, H. B.) (Academic, San Diego, 1991).
- Keller, R. E. & Danilchik, M. *Development* **103**, 193–210 (1988).
- Wilson, P. A. & Keller, R. E. *Development* **105**, 155–166 (1989).
- Mangold, O. *Naturwissenschaften* **21**, 761–766 (1933).
- Hemmati-Brivanlou, A., Stewart, R. M. & Harland, R. M. *Science* **250**, 800–802 (1990).
- Spemann, H. & Mangold, H. *Arch. mikr. Anat. EntwMech.* **100**, 599–638 (1924).
- Nieuwkoop, P. D. *Wilhelm Roux Arch. EntwMech. Org.* **162**, 341–373 (1969).
- Melton, D. A. *Science* **252**, 234–241 (1991).
- Dawid, I. B. in *Methods in Cell Biology* (eds Kay, B. K. & Peng, H. B.) (Academic, San Diego, 1991).
- Slack, J. M. W., Darlington, B. G., Heath, J. K. & Godsave, S. F. *Nature* **326**, 197–200 (1987).
- Kimelman, D. & Kirschner, M. *Cell* **51**, 869–877 (1987).
- Rosa, F. et al. *Science* **239**, 783–785 (1988).
- Smith, J. C. *Development* **99**, 3–14 (1987).
- Asashima, M. et al. *Naturwissenschaften* **77**, 389–391 (1990).
- Sokol, S., Wong, G. G. & Melton, D. A. *Science* **249**, 561–564 (1990).
- McMahon, A. P. & Moon, R. T. *Cell* **58**, 1075–1084 (1989).
- Christian, J. L., Gavin, B. J., McMahon, A. P. & Moon, R. T. *Dev Biol* **143**, 230–234 (1991).
- Sokol, S., Christian, J. L., Moon, R. T. & Melton, D. A. *Cell* **67**, 741–752 (1991).
- Smith, W. C. & Harland, R. M. *Cell* **67**, 753–765 (1991).
- Thomsen, G. et al. *Cell* **63**, 485–493 (1990).
- Asashima, M. et al. *Proc. natn. Acad. Sci. U.S.A.* **88**, 6511–6514 (1991).
- Shiurba, R. A., Jing, N., Sakakura, T. & Godsave, S. F. *Development* **113**, 487–494 (1991).
- Ruiz i Altaba, A. & Melton, D. A. *Trends Genet.* **6**, 57–64 (1990).
- Green, J. B. A. & Smith, J. C. *Nature* **347**, 391–394 (1990).
- Green, J. B., Howes, G., Symes, K., Cooke, J. & Smith, J. C. *Development* **108**, 173–183 (1990).
- Sokol, S. & Melton, D. A. *Nature* **351**, 409–411 (1991).
- Ruiz i Altaba, A. & Jessel, T. *Genes Dev.* **5**, 175–187 (1991).
- Herskowitz, I. *Nature* **329**, 219–222 (1987).
- Amaya, E., Musci, T. J. & Kirschner, M. W. *Cell* **66**, 257–270 (1991).
- Kondo, M. et al. *Biochem. biophys. Res. Commun.* **181**, 684–690 (1991).
- Mathews, L. S., Vale, W. W. & Kintner, C. R. *Science* **255**, 1702–1705 (1992).

- Hemmati-Brivanlou, A., Wright, D. A. & Melton, D. A. (in the press).
- Symes, K. & Smith, J. *Development* **101**, 185–408 (1987).
- Smith, J. C., Price, B. M. J., Green, J. B. A., Weigel, D. & Herrmann, B. G. *Cell* **67**, 79–87 (1991).
- Rosa, F. M. *Cell* **57**, 965–974 (1989).
- Cho, K. W. Y., Blumberg, B., Steinbeissier, H. & De Robertis, E. M. *Cell* **67**, 1111–1120 (1991).
- Mohun, T. J., Brennan, S., Dathan, N., Fairman, S. & Gurdon, J. B. *Nature* **311**, 716–721 (1984).
- Sive, H. L., Hattori, K. & Weintraub, H. *Cell* **58**, 171–180 (1989).
- Bolce, M., Hemmati-Brivanlou, A., Kirschner, P. & Harland, R. (in the press).
- Kintner, C. R. & Melton, D. A. *Development* **99**, 311–325 (1987).
- Smith, J. C., Price, B. M. J., Van Nimmen, K. & Huylebroeck, D. *Nature* **345**, 729–731 (1990).
- Kintner, C. R. & Dodd, J. *Development* **113**, 1495–1506 (1991).
- Sharpe, C. R. & Gurdon, J. B. *Development* **109**, 765–774 (1990).
- Vize, P. D., Hemmati-Brivanlou, A., Harland, R. & Melton, D. A. in *Xenopus laevis: Practical uses in Cell and Molecular Biology* (eds Kay, B. K. & Peng, H. B.) 361–381 (Academic, Florida, 1991).
- Nakamura, T., Sugino, K., Titani, K. & Sugino, H. *J. biol. Chem.* **266**, 19432–19437 (1991).
- Dale, L., Howes, G., Price, B. M. J. & Smith, J. C. *Development* **115**, 573–585 (1992).
- Jones, C. M., Lyons, K. M., Lapan, P. M., Wright, C. V. & Hogan, B. L. M. *Development* **115**, 639–647 (1992).
- Krieg, P. A. & Melton, D. A. *Nucleic Acids Res.* **12**, 7057–7070 (1984).
- Newport, J. & Kirschner, M. *Cell* **30**, 675–686 (1982).
- Nieuwkoop, P. D. & Faber, J. *Normal Table of Xenopus laevis (Daudin)* (North-Holland, Amsterdam, 1967).
- Condie, B. G. & Harland, R. M. *Development* **101**, 93–105 (1987).
- Krieg, P., Varnum, S., Wormington, M. & Melton, D. A. *Dev Biol* **133**, 93–100 (1989).
- Jonas, E. A., Snape, A. M. & Sargent, T. D. *Development* **106**, 399–405 (1989).
- Hemmati-Brivanlou, A. et al. *Development* **110**, 325–330 (1990).
- Harland, R. M. in *Methods in Cell Biology* (eds Kay, B. K. & Peng, H. B.) (Academic, San Diego, 1991).
- Hemmati-Brivanlou, A. & Harland, R. M. *Development* **106**, 611–617 (1989).

ACKNOWLEDGEMENTS. We thank P. Kushner for the notochord antibody, P. Wilson for *Xenopus* brachyury, P. Vize for the Xho-4 probe, C. Kintner for communication of results before publication, M. Bolce for technical advice, and P. Klein and our colleagues for discussion.

Pulsed high-energy γ -rays from the radio pulsar PSR1706–44

D. J. Thompson*, Z. Arzoumanian†, D. L. Bertsch*, K. T. S. Brazier‡, N. D'Amico§, C. E. Fichtel*, J. M. Fierro||, R. C. Hartman*, S. D. Hunter*, S. Johnston¶, G. Kanbach‡, V. M. Kaspi†, D. A. Kniffen#, Y. C. Lin||, A. G. Lyne**, R. N. Manchester¶, J. R. Mattox*††, H. A. Mayer-Hasselwander‡, P. F. Michelson||, C. v. Montigny‡, H. I. Nel‡‡§§, D. Nice†, P. L. Nolan||, K. Pinkau‡, H. Rothermel‡, E. J. Schneid||, M. Sommer‡, P. Sreekumar§§ & J. H. Taylor†

* NASA Goddard Space Flight Center, Greenbelt, Maryland 20771, USA

† Department of Physics, Princeton University, Princeton, New Jersey, 08544, USA

‡ Max-Planck-Institut für Extraterrestrische Physik, D-8046 Garching, Germany

§ Istituto di Fisica dell'Università di Palermo and Istituto di Radioastronomia del CNR, I-40126 Bologna, Italy

|| W. W. Hansen Experimental Physics Laboratory and Department of Physics, Stanford University, Stanford, California 94305 USA

¶ Australia Telescope National Facility, CSIRO, PO Box 76, Epping, New South Wales 2121, Australia

Department of Physics, Hampden-Sydney College, Hampden-Sydney, Virginia 23943, USA

** Department of Physics, University of Manchester, Jodrell Bank, Macclesfield, SK11 9DL, UK

†† Compton Observatory Science Support Center, Computer Sciences Corporation, NASA Goddard Space Flight Center, Greenbelt, Maryland 20771, USA

§§ USRA, NASA Goddard Space Flight Center, Greenbelt, Maryland 20771, USA

|| Grumman Aerospace Corporation, Bethpage, New York 11714, USA

OF the more than 500 known radio pulsars, only three have been detected as γ -ray sources. Using the Energetic Gamma Ray Experiment Telescope (EGRET) on the Compton Gamma Ray Observatory satellite, we have detected pulsed γ -radiation, above 100 MeV in energy, from a fourth radio pulsar, PSR1706–44. Within the pulse period of 102 ms, the γ -emission forms a single broad peak, in contrast to the two narrow peaks seen in the other high-energy γ -ray pulsars. The emission mechanism in all cases is probably the same, the differences arising from the geometry of the magnetic and rotation axes and the line of sight. In these pulsars, γ -ray emission accounts for as much as 1% of the total neutron star spin-down energy, much more than emerges at optical or radio frequencies, so study of this emission is important in understanding pulsar emission and evolution.

Following the SAS-2 and COS-B high-energy γ -ray observations of the Crab (0531+21) and Vela (0833–45) pulsars^{1,2}, searches were made for additional γ -ray pulsars^{3–5}. No other positive results were then found, but calculations have indicated that γ -ray emission is likely from other rotation-powered pulsars^{6,7}. PSR1509–58 has recently been seen at γ -ray energies below 2 MeV by the Burst and Transient Source Experiment (BATSE)⁸ on the Compton Gamma Ray Observatory (GRO), and the high-energy γ -ray source Geminga (2CG195+04) has been found by EGRET to have periodic emission characteristic of an isolated pulsar^{9,10}.

EGRET¹¹ has the conventional components used in high-energy γ -ray telescopes: an anticoincidence system to discriminate against charged particles, a spark chamber system with interspersed pair conversion material to determine the trajec-

tories of the secondary electron-positron pair, a triggering telescope that detects the presence of charged particles with the correct direction of motion, and an energy-measuring device, which in the case of EGRET is a NaI(Tl) crystal. The telescope covers energies from ~ 20 MeV to over 20 GeV with more than an order of magnitude greater sensitivity than the SAS-2 or COS-B instruments, as well as improved resolution in energy, angle and timing measurements^{12,13}. Because of the very low flux level of the high-energy γ -rays, observing periods are typically about 2 weeks.

From 12–26 July, 1991, EGRET viewed a region encompassing the Galactic Centre. Because of its wide field of view, EGRET had good sensitivity in the direction of the γ -ray source identified¹⁴ by COS-B as 2CG342–02, and this source was readily visible in the EGRET data. In addition to this primary observation, EGRET had three limited exposures to this region: 12–27 December, 1991, 19 March–2 April, 1992, and 28 April–7 May, 1992. The source is detectable in all these observations. The radio pulsar PSR1706–44 is positionally coincident with this source. This pulsar was discovered in a search of the southern galactic plane (well after the end of the COS-B mission) using the Parkes radiotelescope at the relatively high frequency of 1,500 MHz (ref. 15). PSR1706–44 has a short period, 102.4 ms, a rapid spin-down rate giving it a characteristic age of only 17,300 years and considerable timing noise (S.J., R.N.M., V.M.K., A.G.L. & N.D'A., manuscript in preparation). The radio pulse is narrow, with a half-power width of 6 ms or $\sim 6\%$ of the period. On the basis of its dispersion measure of $75.68 \text{ cm}^{-3} \text{ pc}$, its distance is ~ 1.5 kpc, about three times the distance to the Vela pulsar. It is therefore plausible that this pulsar should be detectable as a γ -ray source; if its γ -ray luminosity is the same fraction of its total energy-loss rate as for the Vela pulsar, then it should be about a factor of 20 less bright than Vela.

PSR1706–44 is being monitored regularly as part of a coordinated program involving radioastronomers and the Compton GRO instrument teams. Radio timing based on a series of radio observations made at Parkes, overlapping the times of the γ -ray observations, gives the following parameters: frequency $f = 9.7610186961045 \text{ Hz}$, frequency derivative $\dot{f} = -8.86255 \times 10^{-12} \text{ Hz s}^{-1}$, epoch $t_0 = \text{Julian day } 2448658.5$.

For all four observations, γ -rays from the direction of the source were selected within a cone of radius $\theta \leq 5.85^\circ (E_\gamma/100)^{-0.534}$, with energy E_γ in MeV, which reflects the energy-dependent angular resolution of the EGRET instrument¹⁶. Figure 1 shows the arrival times of γ -rays with energy above 500 MeV folded using the specific radio parameters, after correction to the Solar System barycentre. A broad pulse structure, with a width of $\sim 35\%$ of the phase, was seen in the primary

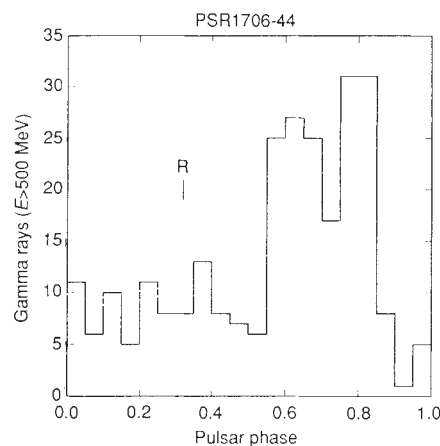


FIG. 1 Gamma-ray phase plot (light curve) for PSR1706–44. Arrival times of EGRET γ -rays above 500 MeV were converted to pulsar phase using the radio period and period derivative. The R in the figure shows the arrival time of the radio pulse relative to the γ -rays.

‡‡ On leave from Department of Physics, Potchefstroom University, Potchefstroom 2520, South Africa.

observation, and the phase plots of the secondary observations are consistent with that distribution in both phase and width. The χ^2 value for the light curve in Fig. 1 is 127 for 19 degrees of freedom, giving a formal probability of chance occurrence of less than 10^{-17} . The H -test¹⁷, an unbinned statistical method, gives a value of 98, which indicates that the formal probability (O. C. DeJager, personal communication) of this being a random distribution is 10^{-17} . Below 500 MeV, the pulsed radiation is harder to recognize above the diffuse galactic radiation, and no pulsed emission is visible below 100 MeV. No statistically significant time structure is seen within the envelope of the pulse, although the shape suggests two wide pulses separated by 0.2 in phase. This pulse shape is markedly different from those of the Crab, Vela and Geminga γ -ray pulsars, all of which show two narrow peaks with a separation of 0.4–0.5 period. Figure 1 also shows the arrival time of the radio pulse, which occurs 0.22 ± 0.02 phase before the leading edge of the γ -ray pulse. This is a larger offset than seen for Vela, where the radio pulse leads the peak of the first γ -ray pulse by 0.126 ± 0.006 phase¹⁹.

Figure 2 shows a preliminary energy spectrum of the pulsed γ -radiation above 100 MeV, based on the primary observation in July 1991. Pulsed emission is seen to at least 5,000 MeV. The spectrum during this fraction 0.35 of the pulse period can be fitted with a power law

$$dN/dE_\gamma = (8.93 \pm 0.76) \times 10^{-10} (E_\gamma/622.4)^{-1.72 \pm 0.08} \quad (1)$$

photons $\text{cm}^{-2} \text{s}^{-1} \text{MeV}^{-1}$

with E_γ in MeV.

The energy scale factor, 622.4 MeV, is chosen so that the statistical errors in the power law index and the overall normalization are uncorrelated. This spectrum is considerably flatter than the pulsed spectrum seen from the Crab (ref. 19; and P.L.N. *et al.*, manuscript in preparation), which has a spectral index ≤ -2.0 . The pulsed emission from Vela, although time-variable, has been seen to have a flat spectrum similar to that of PSR1706–44 during some observations (ref. 20; and G.K. *et al.*, manuscript in preparation).

The time-averaged spectrum has the same form as equation (1), but with a normalization of $(3.12 \pm 0.27) \times 10^{-10}$. The time-

averaged flux of PSR1706–44 above 100 MeV is $(1.0 \pm 0.3) \times 10^{-6}$ photons $\text{cm}^{-2} \text{s}^{-1}$, where the quoted uncertainty includes both the statistics and an uncertainty from the spectral fitting. Between 100 and 5,000 MeV, the energy flux is $(8.2 \pm 2.5) \times 10^{-10}$ erg $\text{cm}^{-2} \text{s}^{-1}$. For a distance of 1.5 kpc and beaming into 3.4 steradian (an assumed value based on the pulse width), this represents 6.1×10^{34} erg s^{-1} or about 1.8% of the spin-down energy loss of 3.4×10^{36} erg s^{-1} , comparable to that seen in the Vela pulsar for similar assumptions²⁰. Except for the shape of the light curve, therefore, the observed features of PSR1706–44 strongly resemble those of the Vela pulsar.

Gamma-ray emission from fast pulsars can be interpreted as radiation (curvature radiation, inverse Compton scattering or synchrotron radiation) from particles accelerated in the pulsar magnetosphere, either near the surface²¹ or in vacuum gaps in the outer magnetosphere²². The broad pulse seen for PSR1706–44 implies that the geometry of the spin axis and magnetic axis relative to the direction of the observer differs from that for the Crab, Vela or Geminga. Future observations may reveal timing structure within the PSR1706–44 pulse which will clarify the geometry. The phase diagram does not preclude the possibility of a double-pulse structure with a narrower separation than those seen for the Crab and Vela γ -ray pulsars. $\square$

Received 27 May; accepted 17 August 1992.

1. Thompson, D. J., Fichtel, C. E., Hartman, R. C., Kniffen, D. A. & Lamb, R. C. *Astrophys. J.* **213**, 252–262 (1977).
2. Kanbach, G. *et al. Astr. Astrophys.* **90**, 163–169 (1980).
3. Ogelman, H. B., Fichtel, C. E., Kniffen, D. A. & Thompson, D. J. *Astrophys. J.* **209**, 584–591 (1976).
4. Thompson, D. J., Bertsch, D. L., Hartman, R. C. & Hunter, S. D. *Astr. Astrophys.* **127**, 220–223 (1983).
5. Buccheri, R. *et al. Astr. Astrophys.* **128**, 245–251 (1983).
6. Harding, A. K. *Astrophys. J.* **245**, 267–273 (1981).
7. Ruderman, M. & Cheng, K. S. *Astrophys. J.* **335**, 306–318 (1988).
8. Wilson, R. B. *et al. IAU Circ. No. 5429* (1992).
9. Halpern, J. P. & Holt, S. S. *Nature* **357**, 222–224 (1992).
10. Bertsch, D. L. *et al. Nature* **357**, 306–307 (1992).
11. Hughes, E. B. *et al. IEEE Trans. Nucl. Sci.* **NS-27**, 364–369 (1980).
12. Kanbach, G. *et al. Space Sci. Rev.* **49**, 69–84 (1988).
13. Kanbach, G. *Gamma Ray Obs. Sci. Workshop Proc.*, 2-1–2-10 (1989).
14. Swanenburg, B. N. *et al. Astrophys. J.* **243**, L69–L73 (1981).
15. Johnston, S. *et al. Mon. Not. R. astr. Soc.* **255**, 401–411 (1992).
16. Thompson, D. J. *et al. Astrophys. J. Suppl.* (in the press).
17. DeJager, O. C., Swanepoel, J. W. H. & Raubenheimer, B. C. *Astr. Astrophys.* **221**, 180–190 (1978).
18. Buccheri, R. *et al. Astr. Astrophys.* **69**, 141–142 (1978).
19. Clear, J. *et al. Astr. Astrophys.* **174**, 85–94 (1987).
20. Grenier, I. A., Hermesen, W. & Clear, J. *Astr. Astrophys.* **204**, 117–132 (1988).
21. Daugherty, J. K. & Harding, A. *Astrophys. J.* **252**, 337–347 (1982).
22. Cheng, K. S., Ho, C. & Ruderman, M. *Astrophys. J.* **300**, 500–521 (1986).

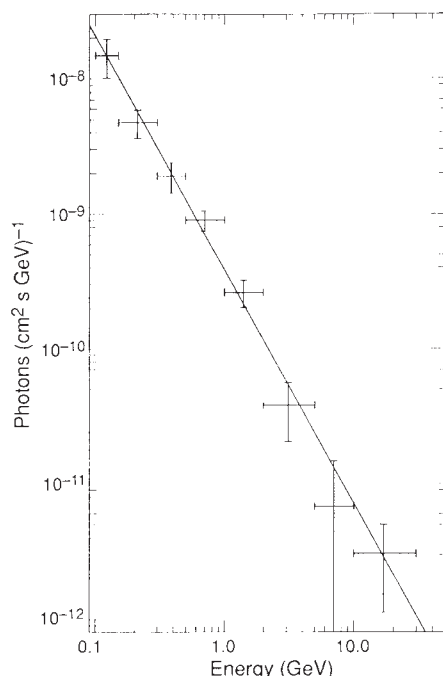


FIG. 2 Energy spectrum of PSR1706–44 during the pulsed phase (0.55–0.90 from Fig. 1). The error bars shown are statistical only.

Pulsar glitches as probes of neutron star interiors

Bennett Link, Richard I. Epstein
& Kenneth A. Van Riper

Los Alamos National Laboratory, Los Alamos, New Mexico 87545, USA

PULSAR rotation rates generally decrease steadily owing to external electromagnetic braking torques, but occasionally show sudden increases ('glitches') followed by gradual recoveries that may last days or years. These events are thought to be consequences of angular momentum transfer between a solid crust, which rotates at the measured pulsar periodicity, and a more rapidly rotating 'loose' component of the neutron star interior. Sudden braking of the differential rotation between the two components will cause a glitch¹, and the subsequent re-establishment of rotational equilibrium between the two components represents the recovery². Earlier studies, using particular models for the coupling between crust and interior, showed that the loose component carries ~2.8% and $\geq 1\%$ of the total moment of inertia of the Vela pulsar³ and PSR 1737–30 (ref. 4) respectively. Here, we analyse post-glitch recovery in four pulsars, and deduce that the loose component

carries at least 0.8% of the total moment of inertia, independent of the form of the coupling. In the context of the 'vortex creep' model of recovery, in which the loose component is the inner-crust neutron superfluid^{2,5-7}, the constraint on the moment of inertia rules out equations of state that are soft at high densities.

Typical pulsar spin behaviour across a glitch is illustrated in Fig. 1. In the absence of glitches or other violent changes, the decelerations of the crust $\dot{\Omega}_c(t) \equiv d\Omega_c(t)/dt$ and the loose component both approach an asymptotic rate $\dot{\Omega}_\infty(t)$. In principle, the loose component can rotate differentially, in which case its angular velocity $\Omega_l(t, r)$ is a function of position. The loose component slows because of drag by the crust, and hence $\Omega_l(t, r) \geq \Omega_c(t)$. When there is no glitch activity, all of the loose component slows and transfers angular momentum to the crust, so $\dot{\Omega}_l(t, r) \leq 0$ and $\dot{J}_l = \int \dot{\Omega}_l(t, r) dI_l \leq 0$, where J_l is the angular momentum of the loose component and I_l its moment of inertia. During a glitch at time t_g (see Fig. 1), the crust increases its angular velocity by an amount $\Delta\Omega_c$, and then decelerates more rapidly than before the glitch. After time t_r , the recovery time, the crust's rotation rate falls below its value before the glitch: $\Omega_c(>t_r) < \Omega_c(<t_g)$. When $t_g < t < t_r$, the crust spins faster than before the glitch, and possibly spins faster than some of the loose component. During this interval, the crust may impart angular momentum to the loose component. After t_r , however, the crust must once again have a lower angular velocity than all of the loose component so that its angular momentum is decreasing: $\dot{J}_l(>t_r) < 0$.

We define $I \equiv I_c + I_l$ to be the total effective moment of inertia of the star, where I_c is the moment of inertia of the solid crust plus any parts of the star tightly coupled to it. From angular momentum conservation,

$$I_c \dot{\Omega}_c(t) + \dot{J}_l(t) = N_{\text{ext}} \quad (1)$$

where the external torque is $N_{\text{ext}} \equiv -I|\dot{\Omega}_\infty(t)|$. Because $\dot{J}_l < 0$ for $t > t_r$ (but before the next glitch) we have

$$I_c |\dot{\Omega}_c(>t_r)| \leq I |\dot{\Omega}_\infty(>t_r)| \quad (2)$$

Although $\dot{\Omega}_\infty(t)$ is not directly observable, it can be compared to the crust spin-down rate. We first note that the external torque decreases as the pulsar evolves so that $|\dot{\Omega}_\infty(>t_r)| < |\dot{\Omega}_\infty(<t_g)|$, although the difference is extremely small. Second, because $\Omega_c(t)$ approaches $\Omega_\infty(t)$ before the glitch, we have $|\dot{\Omega}_\infty(<t_g)| < |\dot{\Omega}_c(<t_g)|$, as can be seen from Fig. 1. With these two inequalities and $I_c = I - I_l$, equation (2) becomes $(I - I_l)|\dot{\Omega}_c(>t_r)| < I|\dot{\Omega}_c(<t_g)|$, and hence

$$\frac{I_l}{I} > 1 - \frac{\dot{\Omega}_c(>t_r)}{\dot{\Omega}_c(<t_g)} \quad (3)$$

A lower limit on I_l can be obtained for any time after t_r and before the next glitch, but the strongest constraint is obtained for times close to t_r , when $|\dot{\Omega}_c|$ is still relatively large. The lower limit on I_l/I derived from equation (3) is less than the fractional change in $\dot{\Omega}_c$ across the glitch, as some of the discontinuity in the spin-down rate decays by t_r .

To account for uncertainties in glitch times, we use the following procedure. We extrapolate the fit of $\Omega_c(t)$ for the pre-glitch data to the time of the first post-glitch measurement at time t'_g ; we denote this quantity by $\Omega_{\text{ex}}(t'_g)$ (see Fig. 1). The corresponding recovery time, t'_r , is found from $\Omega_c(t'_r) = \Omega_{\text{ex}}(t'_g)$. Because t'_r always exceeds t_r , we use $\dot{\Omega}_c(t'_r)$ for $\dot{\Omega}_c(>t_r)$ in equation (3). For the quantity $\dot{\Omega}_c(<t_g)$ we evaluate $\dot{\Omega}_c(t)$ at the time of the last pre-glitch observation. In practice $\dot{\Omega}_c(<t_g)$ varies so slowly that the precise time used is unimportant.

Table 1 shows lower limits on I_l/I obtained from data on selected glitches in PSRs 0531+21 (Crab), 0833-45 (Vela), 0355+54 and 0525+21. We also tabulate the approximate glitch date (which is $\leq t'_g$), the discontinuities in Ω_c and $\dot{\Omega}_c$ across each glitch (indicated with the subscript g), and the interval $t'_r - t'_g$. In the cases of the Crab pulsar and PSR0525+21, recovery

occurs so soon after the glitch (as reflected in the small values for $t'_r - t'_g$) that little of the initial discontinuity in $\dot{\Omega}_c$ decays, and the limit on I_l is $I_l/I \geq (\Delta\dot{\Omega}_c/\dot{\Omega}_c)_g$. In the Vela pulsar, on the other hand, much of the discontinuity in $\dot{\Omega}_c$ decays before t'_r . The 1975 giant glitch of the Vela pulsar yields the most stringent constraint on the moment of inertia of the loose component: $I_l/I \geq 0.0081$.

What is the loose component likely to be? The most promising candidate is the neutron superfluid in the inner crust. This region, where nuclei coexist with a free neutron gas, extends from the density of neutron drip $4.3 \times 10^{11} \text{ g cm}^{-3}$, to the density of nuclear matter, $\rho_0 = 2.8 \times 10^{14} \text{ g cm}^{-3}$. The angular velocity of a superfluid is determined by the spatial arrangement of its vortex lines. In regions of the inner crust, vortex interactions with nuclei inhibit vortex motion so that the superfluid responds only slowly to changes in the angular velocity of the crust^{2,5-7}, and the coupling between this portion of the superfluid and the crust is weak. By contrast, in the core, where superfluid neutrons co-exist with superconducting protons, entrainment of protons by the flowing neutrons induces magnetic moments about the vortices. Electron scattering from magnetized vortices causes the core superfluid to achieve corotation with the crust within minutes after a glitch⁸. Over timescales typical of post-glitch response, therefore, the solid crust and core behave essentially as a single rigid body.

Hereafter we assume that the loose component is the inner-crust superfluid, and write $I_s \equiv I_l$. In principle, the loosely coupled component could comprise all of the free neutrons of the inner crust. Within current theoretical uncertainties, inner-crust superfluidity may extend up to nuclear density⁹ ρ_0 , or may cease to exist above densities as low as $\sim 0.27 \rho_0$ (ref. 10). To allow for the possibility that only a fraction of the free neutrons of the inner crust contributes to I_s , we treat the maximum density to which superfluidity extends as a parameter ρ_{max} .

In general, for a given mass, stellar models with softer equations of state have thinner inner crusts with smaller moments of inertia. In Fig. 2, we show the fractional moment of inertia of the crustal superfluid as a function of neutron star mass for selected maximum densities for superfluidity ($\rho_{\text{max}}/\rho_0 = 1.0, 0.7$ and 0.5) and the following equations of state: BPS (Reid soft-core nucleon potential and hyperons¹¹), P (Reid potential and pure neutron matter¹²), FP (two- and three-nucleon interactions¹³), BJ (Reid potential with short-range repulsion and hyperons¹⁴), CCLR (two-body Levinger-Simmons potential¹⁵) and PS (nucleon coupling to a mean scalar field¹⁶).

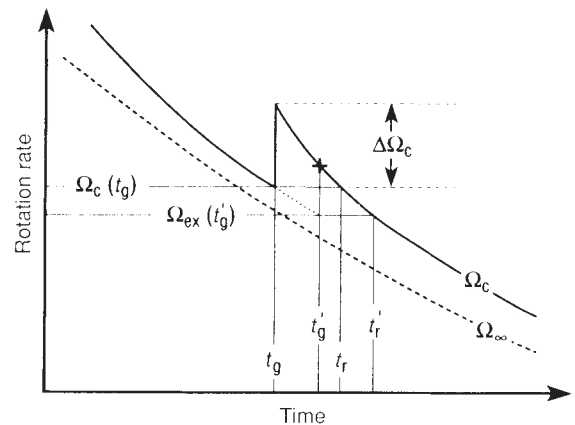


FIG. 1 Schematic evolution of the crust angular velocity Ω_c through a glitch. After a glitch at time t_g , the pulsar recovers its pre-glitch rotation rate at t_r . The first post-glitch measurements (shown by the +) are at t'_g and the corresponding recovery time is t'_r . The dashed line represents the asymptotic spin rate Ω_∞ .

TABLE 1 Glitch parameters

PSR	$\Omega/2\dot{\Omega}$ (yr)	Date (month/yr)	Ref.	$10^8(\Delta\Omega/\Omega)_g$	$10^2(\Delta\dot{\Omega}/\dot{\Omega})_g$	$t'_g - t_g$ (day)	$10^2 I_s/I >$
0531 + 21 (Crab)	1.2 K	2/75	21	3.8	0.25	0.04	0.25
		8/86	22	0.92	0.25	0.008	0.23
0833 - 45 (Vela)	11 K	2/69	23	234	1.0	19	0.63*
		8/71	23	205	1.5	17	0.72
		9/75	23	199	1.1	16	0.81
		7/78	23	306	1.8	25	0.70
		10/81	23	114	0.84	9	0.67
			24	115	4.9	9	0.70
		8/82	23	205	2.3	17	0.65
			24	206	6.3	17	0.74
		7/85	24	160	1.4	13	0.68
		12/88	25	181	7.7	15	0.76
0355 + 54	0.6 M	1/85	26	0.56	0.18	2.3	0.16
0525 + 21	1.5 M	1/74	27	0.13	0.46	1.4	0.37

* In using the fitting formulae of ref. 23, we used the pre-glitch $\dot{\Omega}_c$ extrapolated to t'_g .

Observations suggest that neutron star masses are close to 1.4 solar masses ($M_\odot$). For example, the masses of PSR1913+16 and its companion (also believed to be a neutron star) are $1.442 \pm 0.003 M_\odot$ and $1.386 \pm 0.003 M_\odot$, respectively¹⁷. The BPS equation of state is inconsistent with these observations as it supports a maximum stable mass of only $1.414 M_\odot$, which is $\sim 2\%$ lower than the measured mass of PSR1913+16.

In Fig. 2 our lower limit $I_s/I \geq 0.0081$ is shown. The model-dependent estimate³ of $I_s/I \approx 0.028$ is shown for comparison. We see that for $\rho_{\max}/\rho_0 = 1.0$, the model-independent constraint on I_s is barely consistent with the Vela pulsar being a $1.4 M_\odot$ neutron star with a soft equation of state (such as BPS), and the model-dependent estimate excludes this possibility. For the more plausible case of $\rho_{\max}/\rho_0 = 0.7$, soft equations of state are ruled out for any reasonable mass of the Vela pulsar, whereas a moderately stiff equation of state (such as FP) is consistent with the model-dependent estimate. For superfluidity extending to $0.5 \rho_0$, moderately stiff equations of state are allowed by the model-independent constraint, although not by the model-dependent estimate (BPS, P and FP all excluded).

More recent equations of state are stiffer than BPS and P, and yield maximum neutron star masses of $1.8\text{--}2.2 M_\odot$ (see for example, ref. 18). Brown¹⁹ has argued, however, that the equation of state may be much softer than generally thought, and suggests that the maximum stable neutron star mass is near $1.5 M_\odot$. Brown *et al.*²⁰ propose that the failure to observe a pulsar in supernova 1987A indicates that the core collapsed to form a black hole, as would be the case if the equation of state were very soft. The theoretical situation is far from resolved, but the method presented here, when applied to future timing data, may help settle this issue. $\square$

Received 3 April; accepted 7 August 1992.

- Anderson, P. W. & Itoh, N. *Nature* **256**, 25-27 (1975).
- Alpar, M. A., Anderson, P. W., Pines, D. & Shaham, J. *Astrophys. J.* **276**, 325-334 (1984).
- Pines, D. & Alpar, M. A. in *Structure and Evolution of Neutron Stars* (eds Pines, D., Tamagaki, R. & Tsuruta, S.) (New York, Addison-Wesley, in the press).
- Michel, F. C., Bland Hawthorn, J. & Lyne, A. G. *Mon. Not. R. astr. Soc.* **246**, 624-627 (1990).
- Link, B. & Epstein, R. I. *Astrophys. J.* **373**, 592-603 (1991).
- Link, B., Epstein, R. I. & Baym, G. *Astrophys. J.* (in the press).
- Epstein, R. I. & Baym, G. *Astrophys. J.* **328**, 680-690 (1988).
- Alpar, M. A., Langer, S. A. & Sauls, J. A. *Astrophys. J.* **282**, 533-541 (1985).
- Ainsworth, T. L., Wambach, J. & Pines, D. *Phys. Lett. B* **222**, 173-178 (1989).
- Chen, J. M. C., Clark, J. W., Krotscheck, E. & Smith, R. A. *Nucl. Phys. A* **451**, 509-540 (1986).
- Baym, G., Pethick, C. & Sutherland, P. *Astrophys. J.* **170**, 299-317 (1971).
- Pandharipande, V. R. *Nucl. Phys. A* **174**, 641-656 (1971).
- Friedman, B. & Pandharipande, V. R. *Nucl. Phys. A* **361**, 502-520 (1981).
- Bethe, H. A. & Johnson M. *Nucl. Phys. A* **230**, 1-58 (1976).
- Cohen, J. M., Langer, W. D., Rosen, L. C. & Cameron, A. G. W. *Astrophys. Space Sci.* **6**, 228-239 (1970).
- Pandharipande, V. R. & Smith, R. A. *Phys. Lett. B* **59**, 15-18 (1975).
- Wolszczan, A. *Nature* **350**, 688-690 (1991).
- Wiringa, R. B., Fiks, V. & Fabrocini, A. *Phys. Rev. C* **38**, 1010-1037 (1988).
- Brown, G. E. *Nature* **336**, 519-520 (1988).
- Brown, G. E., Bruenn, S. W. & Wheeler, J. C. *Comments Astrophys.* (in the press).
- Lohsen, E. H. G. *Astr. Astrophys. Suppl.* **44**, 1-14 (1981).
- Lyne, A. G. & Pritchard, R. S. *Mon. Not. R. astr. Soc.* **229**, 223-226 (1987).
- Cordes, J. M., Downs, G. S. & Krause-Polstorff, J. *Astrophys. J.* **330**, 847-869 (1988).
- McCulloch, P. M., Klekociuk, A. R., Hamilton, P. A. & Royle, G. W. R. *Aust. J. Phys.* **40**, 725-730 (1987).
- McCulloch, P. M., Hamilton, P. A., McConnell, D. & King, E. A. *Nature* **346**, 822-824 (1990).
- Lyne, A. G. *Nature* **326**, 569-571 (1987).
- Downs, G. S. *Astrophys. J.* **257**, L67-L70 (1982).

ACKNOWLEDGEMENTS. This work was carried out under the auspices of the U.S. Department of Energy.

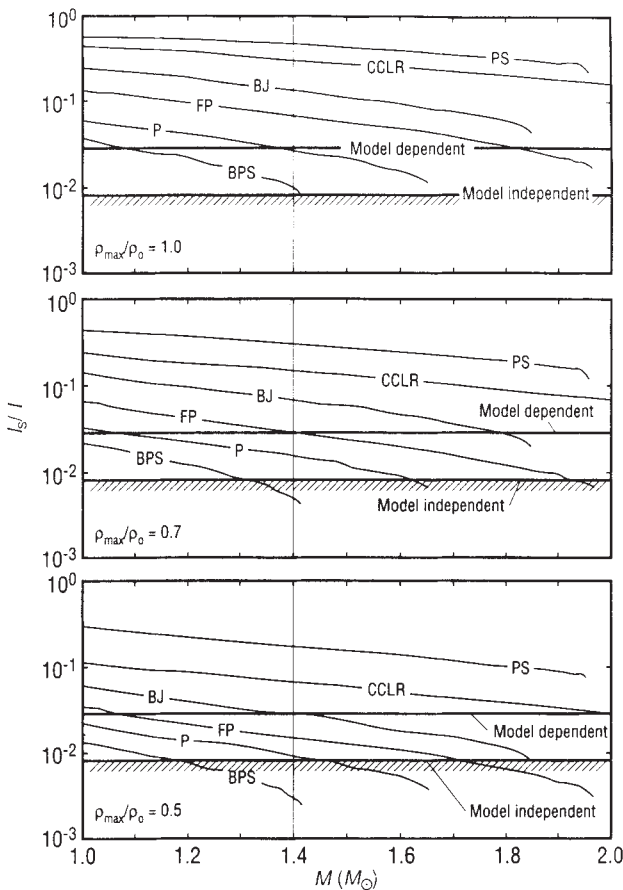


FIG. 2 Fractional moment of inertia of the inner crust superfluid as a function of neutron star mass for selected equations of state and maximum densities $\rho_{\max}$ for superfluidity. The model-independent constraint obtained here and the model-dependent estimate obtained in ref. 3 are shown. Values of I_s that fall in the cross-hatched regions are ruled out.

Gravitational wakes in Saturn's rings

H. Salo

Department of Astronomy, University of Oulu, 90570 Oulu, Finland

THE outer parts of Saturn's rings display a variety of local non-uniformities in their particle distributions. Azimuthal brightness variations are seen in the A-ring^{1,2}, and may be attributable to the gravitational aggregation of particles into linear wakes that trail the rotation of the ring³⁻⁵; Voyager's stellar occultation experiments revealed large differences, on length scales as small as 150 m, in the surface density of particles⁶. Theoretical arguments^{7,8} suggest that local instabilities may occur in both the A- and B-rings, the instability criterion depending on the velocity dispersion of ring particles and the orbital velocity and mass density in the ring. These arguments, however, are derived from the purely gravitational dynamics of an idealized, infinitesimally thin ring, made of identical particles. Here I use numerical simulations, including both gravitational interactions and dissipative impacts between particles, to study realistic models of Saturn's rings. For the C-ring there is no instability, but for the B- and A-rings gravitational wakes form. In the A-ring these wakes are so strong that particles trapped in them form metre-sized aggregate particles, which themselves lead to further instability. These different behaviours are consistent with the observational evidence.

The dynamical evolution of the main components of Saturn's rings is governed by the frequent impacts among the macroscopic icy particles; the local equilibrium state follows from the balance between dissipative energy loss in partially inelastic collisions and viscous gain of energy from the systematic velocity field. Laboratory measurements⁹ indicate that collisions are very inelastic, with coefficient of restitution $\varepsilon < 0.5$, for impact velocities of only a few millimetres per second. Accordingly, purely collisional simulations predict little velocity dispersion at equilibrium, and therefore predict that the vertical thickness of the rings should only be only few times the dominant particle size. It is instructive to apply gravitational stability criteria to such idealized systems. According to Toomre¹⁰, a differentially rotating, infinitely flat system is stable against all axisymmetric perturbations, as long as the radial velocity dispersion exceeds $\sigma_r > \sigma_{cr} = 3.36G\Sigma/\kappa$, where Σ denotes the surface mass density and κ the epicyclic frequency, equal to angular frequency Ω in a keplerian velocity field. For a system of identical particles each of density ρ and radius r , the surface density is related to

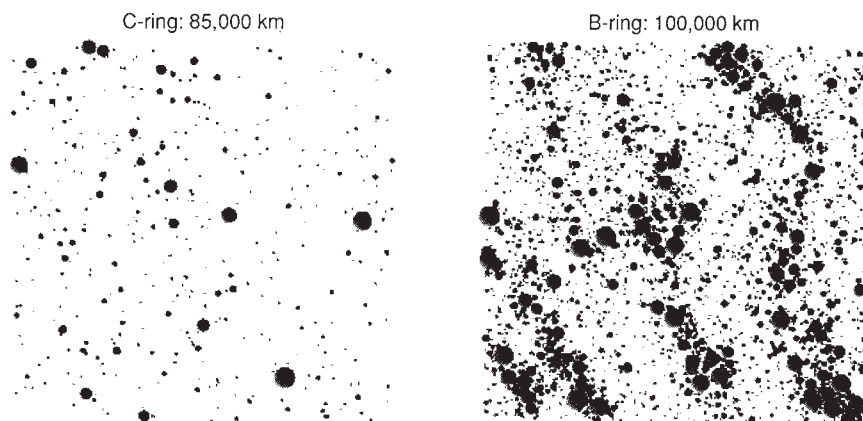
normal optical thickness τ by $\Sigma = 4\rho r\tau/3$. Collisional simulations^{11,12} indicate that $\sigma_r \approx 3r\Omega$ or smaller for any τ , and the stability criterion can thus be written $\tau_{cr} = 0.4(10^8/a \text{ m})^3 (0.9/\rho \text{ g cm}^{-3})$, giving the value of τ above which instability should occur at a distance a from the centre of Saturn. For the C-, B- and A-rings, $\tau_{cr} \approx 0.7, 0.4$ and 0.2 , respectively, suggesting that although the C-ring, with mean $\tau \approx 0.15$, should be stable, both B- and A-rings ($\tau > 1.0$ and $\tau \approx 0.5$) might be susceptible to local instabilities.

This simple estimate for τ_{cr} should be considered with caution, as it ignores the increase of random velocities due to gravitational encounters¹³⁻¹⁵. In addition, recent collisional simulations with extended size distribution¹⁶ indicate that σ_r can be as much as 5-fold for the smallest particles as compared with σ_r of the largest ones. This might stabilize at least their subpopulation, and possibly the whole system, by decreasing the effective surface density susceptible to perturbations. Any wakes formed would also scatter particles, and the net result might well be to alter the velocity dispersion in such a way that instabilities are eventually avoided¹⁰.

Even if the axisymmetric Toomre stability criterion is fulfilled, systems with the ratio $Q_T = \sigma_r/\sigma_{cr}$ close to unity can still show a strong tendency to form transient density wakes. According to Julian and Toomre¹⁷, these wakes appear in the form of trailing wavelets whose most prominent azimuthal wavelength is $\sim 2\lambda_{cr}$, where $\lambda_{cr} = 4\pi^2 G\Sigma/\kappa^2$ denotes the shortest axisymmetric wavelength stabilized by differential rotation alone. For Saturn's rings, $\lambda_{cr} \approx 70(a/10^8)^3(\Sigma/1,000)$, where a is in metres and Σ in kg m^{-2} , or about 6, 65 and 85 m for typical C-, B- and A-ring parameters. The expected pitch angle (angle between wavecrest and tangential direction) is $\sim 10\text{--}15^\circ$ for a keplerian velocity field, if the thickness of the system is negligible compared with λ_{cr} . At least for stellar systems, unless there is some external driving (for example in the form of an orbiting mass concentration), these wakes are rapidly damped¹⁷. Therefore, the formation of particle aggregates seems to be necessary to maintain persistent wakes. Formally, the classical Roche distance for ice particles $R_{\text{Roche}} = 136,500 \text{ km}$, but Weidenshilling *et al.*¹⁸ estimate that particle aggregates could form well inside R_{Roche} : in principle, a small particle can stick gravitationally to the surface of a slowly rotating large ice particle for $a > 70,000 \text{ km}$, or everywhere in the rings. As this estimate does not take into account the disrupting effects of collisions, however, the survival of particle aggregates is uncertain. On the other hand, external driving may not be necessary in the presence of dissipation: impacts cool the system continuously and thereby aid the amplification of new wakes (compare with stellar

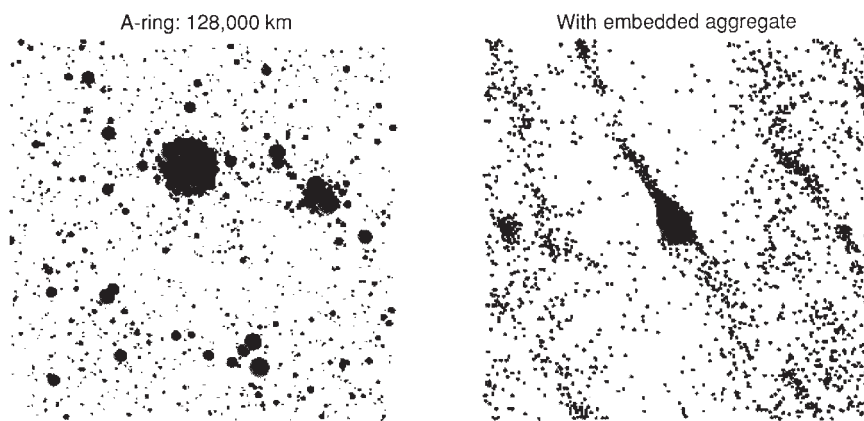
FIG. 1 Simulation for Saturn's C- and B-rings, for

distances of $a = 85,000$ and $100,000 \text{ km}$ from the centre of Saturn. I use the local simulation method¹⁶, analogous to the stellar dynamical simulations of Toomre²⁰ and the planetary ring simulations of Wisdom and Tremaine¹¹: all the calculations are restricted to a small co-moving cell inside the rings, with periodic boundary conditions taking into account the systematic shearing motion. Particle distributions after 10 orbital revolutions are displayed. The size distribution is approximated by a power-law with slope 3, extending over $50 \text{ cm} < r < 5 \text{ m}$, and each side of the (square) calculation cell is 170 m. Direction of the planet is to the left and mean orbital motion is upward. The gravitational force on each particle is calculated from all the other particles within 85 m. No softening is used, as the finite size of particles makes it unnecessary. Internal density of particles is taken to be $\rho = 0.9 \text{ g cm}^{-3}$, and the surface densities are 120 and 960 kg m^{-2} , obtained by varying the total number of particles (400 and 3,200). For the coefficient of restitution, I assume a velocity-dependent elasticity model⁹: $\varepsilon(v_p) = (v_p/v_0)^{-0.234}$, $0.25 < \varepsilon < 1$, where v_0 denotes the perpendicular component of impact velocity and $v_0 = 0.01 \text{ cm s}^{-1}$. Initially, random velocities correspond to equilibrium



values in the non-gravitating case. The independence of the results from the cell size was checked by additional runs. The B-ring experiment required about 25 CPU hours on a IBM ES/9121-260 VF mainframe. Particle plots were produced by IDL software.

FIG. 2 Simulation model for Saturn's A-ring. The simulation on the left corresponds to Fig. 1, except that $a=128,000$ km and $\Sigma=480$ kg m $^{-2}$, obtained for 1,600 particles. The system is displayed after 10 orbital revolutions, when initial wakes have disappeared and $\sim 30\%$ of the mass has merged into single aggregate. On the right, the formed particle aggregate has been replaced by a single underdense particle with $r=15$ m and $\rho=0.3$ g cm $^{-3}$, embedded among 1,600 small particles with solid ice density and radius $r=2$ m (total $\Sigma \approx 450$ kg m $^{-2}$). This reduces the computational burden considerably, as most of the CPU time in the previous experiment was spent in calculating the internal collisions among the particles constituting the aggregate. Width of the simulation cell is doubled (340 m) and the figure shows a snapshot of the system after 5 orbital revolutions.



dynamical simulations including cooling¹⁹, where persistent spirals are maintained).

Because so many factors affect the dynamical state of rings, direct particle simulations offer the most reliable method for assessing the gravitational stability and possible presence of gravitational wakes. Figures 1 and 2 show gravitational simulations for three Saturnocentric distances corresponding to the C-, B- and A-rings. The elastic properties of particles follow laboratory measurements⁹, and the size distribution is represented by a power law with slope 3, in accordance with Voyager measurements²¹. Although the distribution is truncated at the lower end, both the maximum particle size, which governs the velocity dispersion through gravitational encounters, and the local surface density, governing collective phenomena, should be fairly representative of those in the actual ring system (see ref. 22).

Figure 1 shows no trace of collective phenomena for the C-ring, because of its low surface density and because it is close to the planet. For the B-ring model, clear wakes are present. Inspection of the simulation system at various times shows that although individual wakes are transient phenomena and lose their identity in a few orbital periods, the overall density variations stay at roughly constant. The scale of these wakes seems to be consistent with the Julian-Toomre estimate, although the pitch angle is about twice that expected for a strictly two-dimensional case, or $\sim 30^\circ$ instead of 15° . The velocity dispersion is roughly twice that seen in a similar run without self-gravity: for the largest particles the observed σ_r corresponds to $Q_T \approx 2$, and for the smallest, $Q_T \approx 3.5$. When gravity is included, the velocity dispersion shows fluctuations of $\sim 30\%$, following the formation and disappearance of individual strong wakes.

For A-ring parameters, the behaviour is different. As for the B-ring, the initially cool system forms strong wakes within the first orbital revolution, but these wakes tend to collapse into elongated, almost radial particle groups, which may merge into a single aggregate (Fig. 2; aggregate size ~ 15 m). This behaviour confirms the proposed¹⁸ rapid formation of aggregates, although it occurs only at the outer portions of the rings. In the case displayed, the aggregate is full of voids, and its final effective density is 30% of that of its constituent particles. Because of the lowered density, additional particles can no longer accumulate at its surface. In several additional runs, including particle spins and friction, and in models including more-elastic impacts, aggregates similar to that of Fig. 2 start forming at $a=125,000$ – $130,000$ km. Because the number of simulation particles is limited, it is difficult to estimate the actual maximum size possible if new particles were continuously available, and if the mutual collisions between aggregates were also included. Including centimetre-sized particles might also aid further growth as they could increase the mean density of the aggregates by filling the empty spaces between larger particles.

Figure 2 (right) displays another run for the A-ring, where the aggregate of the previous experiment has been treated as a single underdense particle. The size of the calculation cell has been doubled, and the wakes excited by the aggregate become visible. Also evident is the zone partially cleared by the massive particle. Compared with the previous B-ring model or the initial A-ring model, the pitch angle of wakes is reduced, especially at the inner edge of the gap.

These dynamical models are consistent with several observations. For example, the distinctly non-Poisson character of the photon noise in the Voyager stellar occultation measurement (PPS) indicates⁶ that for the A-ring, the surface area locally covered by particles deviates significantly between successive measurement points. For the C-ring, uniform density is implied. The extra noise was explained by large maximum particle size, the variance arising from the presence or lack of large particle in each sampling area. According to our experiments, particles with size exceeding 10 m could be identified with the particle aggregates, without any need to revise the cut-off radius of the actual underlying particle distribution. The same explanation was also mentioned in ref. 6.

Dones and Porco⁵ (see also Esposito²³) have shown by light-scattering simulations that Julian-Toomre type wakes can in principle reproduce the observed azimuthal variations in A-ring brightness. Our model of wakes excited by an embedded aggregate (Fig. 2) should therefore be at least qualitatively consistent with the observations. Detailed photometric calculations would be needed to ascertain whether the present model correctly explains the brightness maximum at 165° from the superior conjunction. An unexpected new finding is the predicted wake structure in the B-ring, as no asymmetry has been measured. This could be due to the large optical depth: for example, Franklin and Colombo²⁴ note that even 10% variations in the surface density would not be detectable for this ring. Another possibility²³ is that the wakes are on a smaller scale than for the A-ring. Possible observational support for B-ring wakes is provided by the PPS measurement: according to Showalter and Nicholson⁶, in these regions of B-ring that are not totally opaque, the implied variance is larger than elsewhere in the rings, suggesting structures on a scale of 20 m which would be roughly consistent with the simulated wakes. Dynamical arguments alone strongly suggest that B-ring wakes exist, unless the adopted standard models for $N(r)$, ϵ , Σ or ρ need serious revision. $\square$

Received 8 June; accepted 17 August 1992.

1. Camichel, H. *Ann. Astrophys.* **21**, 231–242 (1958).
2. Thompson, W. K., Lumme, K., Irvine, W., Baum, W. & Esposito, L. *Icarus* **46**, 187–200 (1981).
3. Colombo, G., Goldreich, P. & Harris, A. W. *Nature* **264**, 344–345 (1978).
4. Franklin, F. A. *et al. Icarus* **69**, 280–296 (1987).
5. Dones, L. & Porco, C. *Bull. Am. astr. Soc.* **21** No. 3, 929 (1989).

6. Showalter, M. R. & Nicholson, P. D. *Icarus* **87**, 285–306 (1990).
7. Goldreich, P. & Tremaine, S. A. *Rev. Astr. Astrophys.* **20**, 249–283 (1982).
8. Salo, H. *Earth Moon Planets* **30**, 113–128 (1984).
9. Bridges, F. G., Hatzes, A. P. & Lin, D. N. C. *Nature* **309**, 333–335 (1984).
10. Toomre, A. *Astrophys. J.* **139**, 1217–1238 (1964).
11. Wisdom, J. & Tremaine, S. *Astr. J.* **95**, 925–940 (1988).
12. Salo, H. *Icarus* **90**, 254–270 (1991).
13. Cuzzi, J. N., Burns, J. A., Durisen, R. H. & Hamill, P. *Nature* **281**, 202–240 (1979).
14. Salo, H. *Earth Moon Planets* **33**, 189–200 (1985).
15. Salo, H. *Earth Moon Planets* **38**, 149–181 (1987).
16. Salo, H. *Icarus* **96**, 85–106 (1992).
17. Julian, W. H. & Toomre, A. *Astrophys. J.* **146**, 810–830 (1966).
18. Weidenschilling, S. J., Chapman, C. R., Davis, D. R. & Greenberg, R. in *Planetary Rings* (eds Greenberg, R. & Brahic, A.) 367–415 (Univ. of Arizona Press, Tucson, 1984).
19. Sellwood, J. A. & Carlberg, R. G. *Astrophys. J.* **282**, 61–74 (1984).
20. Toomre, A. in *Dynamics and Interactions of Galaxies* (ed. Wielen, R.) 292–303 (Springer, Berlin, 1990).
21. Zebker, H. A., Marouf, E. A. & Tyler, G. L. *Icarus* **64**, 531–548 (1985).
22. Cuzzi, J. N. *et al.* in *Planetary Rings* (eds Greenberg, R. & Brahic, A.) 73–199 (Univ. of Arizona Press, Tucson, 1984).
23. Esposito, L. W. *Bull. Am. astr. Soc.*, **10** No. 3, 584–585 (1978).
24. Franklin, F. A. & Colombo, G. *Icarus* **33**, 279–287 (1978).

Ferroelectric liquid crystals from achiral molecules

F. Tournilhac*, L. M. Blinov†, J. Simon* & S. V. Yablonsky†

*ESPCI—Chimie et Electrochimie des Matériaux Moléculaires, CNRS-URA 429, 75231 Paris Cedex 05, France

†Institute of Crystallography, Russian Academy of Science, 117333 Moscow, Russia

THE electrical and optical properties of ferroelectric liquid crystals make them of considerable technological interest¹. Although the symmetry of most liquid crystalline phases is too high to allow spontaneous polarization, it has long been recognized that a tilted smectic phase made up of chiral molecules can be ferroelectric, owing to a reduction in the overall symmetry of the material². This forms the basis for conventional ferroelectric liquid crystals, in which the bulk polarization lies within the plane of the smectic layers. Here we describe a new class of organic ferroelectrics composed of achiral 'polyphilic' molecules^{3,4}. Ferroelectricity is demonstrated by measurements of an acoustically induced piezoelectric response and the determination of repolarization currents. The mechanism by which the polar phase is generated differs from that found in conventional chiral smectics, and should in principle allow the preparation of a phase in which the bulk polarization is parallel to the long axis of the constituent molecules.

The physical properties of polar materials are used in applications such as pyroelectric infrared imaging, piezoelectric trans-

ducers and nonlinear optics. The symmetry of most liquid crystalline phases, however, is too high to allow the spontaneous polarization necessary for ferroelectric behaviour. Individual molecules may themselves be strongly polar, but because of the large volume of molecules, the dipole-dipole interaction is usually too weak to produce long-range polar ordering⁵. Another type of interaction is necessary to build up a polar mesophase. In the smectic C* phase, chiral molecules are used to lower the overall symmetry of the material². The strategy for producing a polar axis in this case can be decomposed into two steps (Fig. 1a). First, tilting transforms the $D_{\infty h}$ symmetry of a smectic A phase into a C_{2h} symmetry; second, in a tilted phase composed of one isolated enantiomer, the chirality no longer allows the presence of a mirror plane and the symmetry group of the condensed phase becomes C_2 . This phase is ferroelectric with a polar axis in the plane of the layers. The dipole components parallel to the long molecular axis, if present, are still randomly oriented up and down.

A chemical approach to construct new polar materials has been suggested^{3,4} whereby a 'polyphilic' effect is used to induce the formation of polar self-assemblies. Polyphilic molecules comprise three or more chemically different subunits which tend to segregate into homogeneous microdomains (intralamellar polyphilic effect; Fig. 1b). Moreover, if a molecule is noncentrosymmetric but both its ends have the same chemical nature, this can induce either a centrosymmetric (antiferroelectric) or noncentrosymmetric (ferroelectric) packing of the layers (interlamellar polyphilic effect); longitudinal polarization should therefore be possible in smectic mesophases of polyphilic molecules where the noncentrosymmetric arrangement is electrostatically favoured. Such a mesophase would be of practical

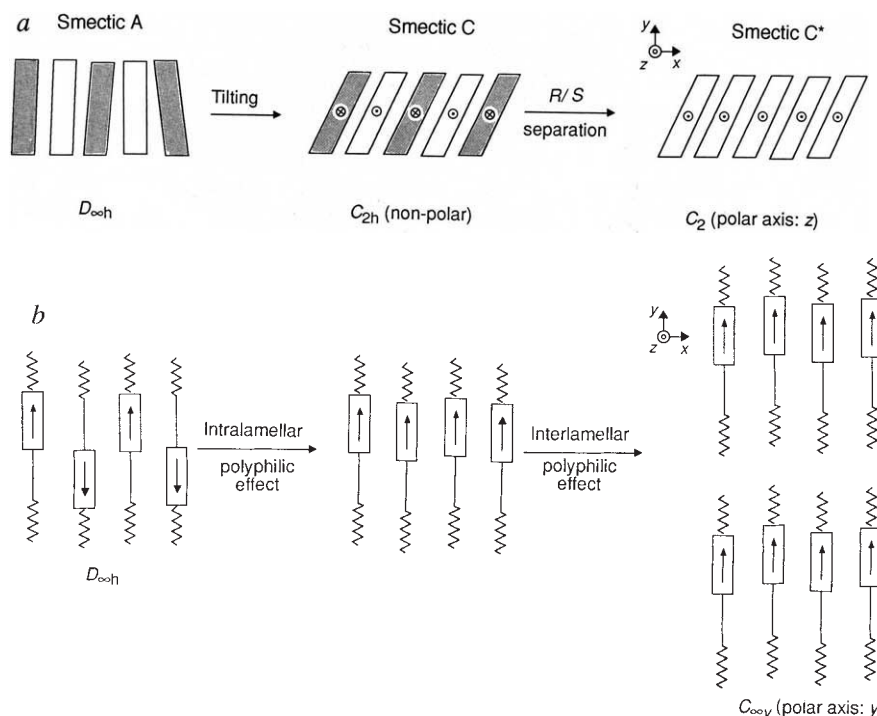


FIG. 1 Two strategies for obtaining polar mesophases. a, Transformation of smectic A to smectic C and selection of one of the chiral enantiomers (shaded/unshaded). b, Intralamellar polyphilic effect: formation of polar smectic layers. Interlamellar polyphilic effect: formation of a polar condensed phase.

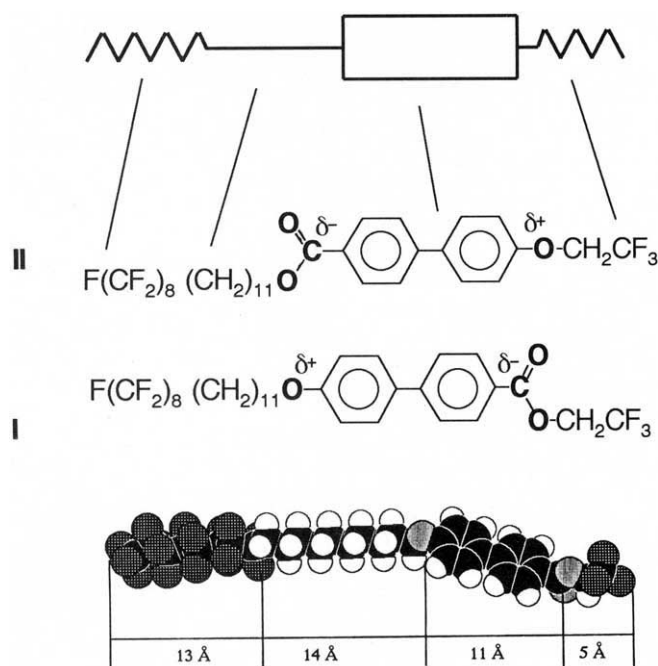


FIG. 2 The two polyphilic compounds investigated. The characteristic dimensions of the various fragments are indicated.

interest in nonlinear optics for which a large charge transfer is required along the long molecular axis⁶.

The possibility of obtaining longitudinal ferroelectric or antiferroelectric liquid crystals has been discussed theoretically^{7,8}. Here we present experimental confirmation of the ferroelectric behaviour of polyphilic compounds. We have studied two derivatives (I and II) (Fig. 2) and their binary mixtures, showing smectic phases referred to as smectic X and X'. The two compounds are made up of three chemically distinct parts: perfluoroalkyl and alkyl side chains, and a diphenyl rigid core. Compounds I and II differ from each other in the direction of their longitudinal dipoles, as the donor and acceptor groups have been exchanged. A mixture of I and II is therefore expected to decrease the intralamellar electrostatic repulsion. Both pure compounds demonstrate a smectic A phase and a metastable and partially ordered smectic X phase. Their binary mixtures yield a stable enantiotropic X' phase with an optical texture slightly different from that of the X phase of the pure compounds. The X and X' phases are miscible in each other and distinct from the crystalline phase. The X-ray diffraction pattern of an unoriented sample of the mixture M70 (70% of I and 30% of II) in its X' phase shows three narrow reflections with Bragg spacings in the ratio 1:2:3, indicating a lamellar organization (interlayer distance 33 Å). At wide angles, no peak is detected other than a weak diffuse ring with a maximum at ~ 4.5 Å and two rather sharp lines of comparable intensities at 4.8 Å and 4.4 Å. These two peaks are probably related to the alkyl and perfluoroalkyl chains, respectively.

We studied two fundamental properties characteristic of ferroelectric substances: the piezoelectric response and the repolarization current. The orientation of dipoles in a material results in apparent charges at the boundaries of the medium. When the material is placed between conductive electrodes, these charges are compensated by carriers of opposite sign at the surface of the electrodes. A change of the overall dipole orientation in the material will therefore correspond to a current (known as repolarization current) through the external circuit. The macroscopic polarization may be calculated by integrating this current.

The polarization can be varied by applying an acoustic stress, which generates a voltage change across the sample (piezoelectric response). The equipment necessary to measure the piezoelectric response has been described previously⁹. In experiments with the pure compound I, the polar state can only be induced using an external d.c. voltage (poling voltage). If the substance is poled during cooling from the isotropic phase, a piezoelectric response stable for days is observed in the metastable smectic X phase. Its amplitude ($U_p \approx 1$ –10 mV) is comparable with the signal observed for conventional chiral smectic C ferroelectrics ($U_p \approx 15$ –20 mV for a ferroelectric with a spontaneous polarization of ~ 100 nC cm⁻²). Nonferroelectric compounds show background piezoelectric signal at the level of a few microvolts.

On heating, the piezovoltage associated with the smectic X phase of I decreases. If crystallization occurs, the signal drops. A complete cycle (poling voltages positive and then negative) can be achieved at 75 °C (Fig. 3). The characteristic time of the repolarization process at this temperature is a few minutes. This precludes the observation of repolarization currents for I.

Repolarization currents could be measured for the mixture M70 (70% of I and 30% of II). In this case, the smectic X' phase is stable over a wide temperature range including room temperature; the macroscopic polarization can, as previously, be induced by an external field. The same hysteresis behaviour is then observed. Moreover, on cooling a sample from the smectic A phase, even if no electric field is applied, spontaneous transition into a polar state occurs within the range 55–30 °C.

The appearance of the polar state is indicated by sharp steps (either one or several, up to five) in the piezoelectric voltage accompanied by narrow peaks in the repolarization current (Fig. 4). The optical textures of the X' phase before and after repolarization are identical. The sign of the current pulses correlates with the sign of the corresponding piezoelectric voltage steps. The

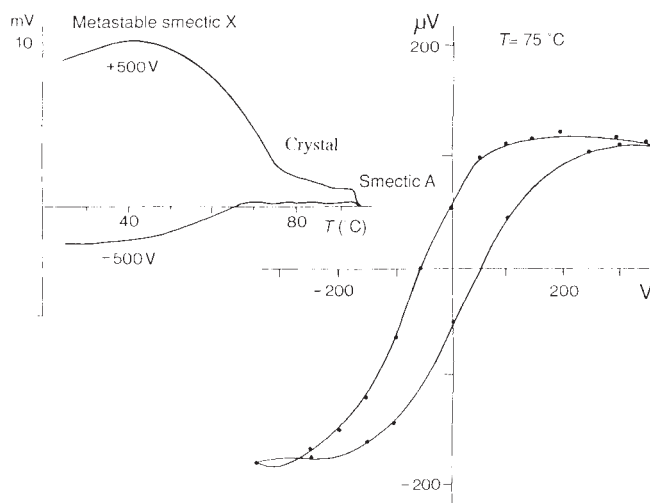


FIG. 3 Hysteresis loop for the piezoelectric signal obtained at 75 °C for the metastable X phase of I. A drop of a substance is placed in a flat capillary (thickness 100 μm) formed by two glass plates coated with optically transparent SnO₂ to form electrodes. Piezoelectricity is induced by mechanical distortion of the material caused by alternating air flow generated by a loudspeaker. The cell is placed in a thermostat and the optical texture of the substance can be observed using polarized-light microscopy. Each point of the hysteresis loop is taken after applying the corresponding d.c. voltage for 2 min. In the left upper corner, the temperature dependences of the piezoelectric response for the same metastable X-phase of I preliminarily poled either with +500 V or -500 V (on cooling down from the isotropic phase) is shown.

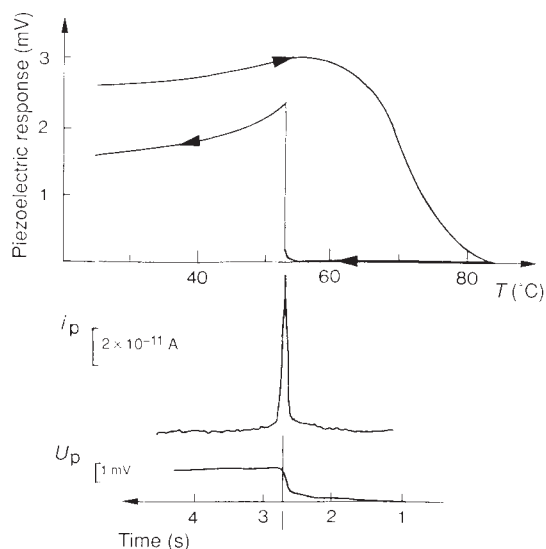


FIG. 4 The temperature behaviour of the piezoelectric response during heating the spontaneously polarized sample of M70 and subsequent cooling cycle. The spontaneous transition to the polar state occurs in the smectic X' phase at $T=53^\circ\text{C}$. The associated jump of the piezoresponse $U_p(t)$ and the corresponding pulse of the repolarization current $i(t)$ flowing through the electrometric amplifier ('fast feedback' mode) are shown below.

spontaneous polarization P_s at the transition is calculated to be in the range $1\text{--}1.5\text{ nC cm}^{-2}$ for the mixture M70 (for U_p signals of $2\text{--}3\text{ mV}$). At room temperature, spontaneously repolarized samples show a stable piezoelectric response.

We have determined other properties characteristic of ferroelectric materials. First, optical second-harmonic generation¹⁰ ($1.06\text{ }\mu\text{m}$) has been observed on unoriented samples of M70. Second, pyroelectric measurement¹¹ on I allowed us to calculate its macroscopic polarization ($\sim 4\text{ nC cm}^{-2}$).

We can thus conclude that the polyphilic compounds investigated show ferroelectric properties. The smectic X' phase consists of polar domains which can spontaneously realign. The direction of the macroscopic polarization can be switched, both in the X and X' phases, by an external field, with hysteresis loops typical of ferroelectrics. No electro-optical effect has been observed during the spontaneous transition; this means that macroscopic polarizations of opposite signs correspond to the same refractive index ellipsoid, as should be the case for longitudinal ferroelectrics. But no firm conclusion can yet be drawn about the mechanism involved in the ferroelectric behaviour of these new materials. The combination of the polyphilic character with longitudinal and transverse molecular dipole moments could produce a variety of ferroelectric and antiferroelectric packings. Boundary effects and layer curvatures might also contribute to the behaviour. □

Received 29 January; accepted 5 August 1992.

- Goodby, J. W. et al. (eds) *Ferroelectric Liquid Crystals* (Gordon and Breach, Philadelphia, 1991).
- Meyer, R. B., Liebert, L., Strzelecki, L. & Keller, P. *J. Phys. (Paris)* **36**, L69–L71 (1975).
- Tournilhac, F., Bosio, L., Nicoud, J. F. & Simon, J. *Chem. Phys. Lett.* **145**, 452–454 (1988).
- Tournilhac, F. & Simon, J. *Ferroelectrics* **114**, 283–287 (1991).
- Beresnev, L. A., Blinov, L. M., Osipov, M. A. & Pikin, S. A. *Molec. Cryst. Liq. Cryst.* **A158**, 1–150 (1988).
- Walba, D. M. et al. *J. Am. chem. Soc.* **113**, 5471 (1991).
- Prost, J. & Barois, P. *J. chim. Phys.* **80**, 65–80 (1983).
- Petschek, R. G. & Wiefing, K. M. *Phys. Rev. Lett.* **59**, 343–346 (1987).
- Blinov, L. M., Davidyan, S. A., Petrov, A. G., Todorov, A. T. & Yablonsky, S. V. *JETP Lett.* (in Russian) **48**, 259–262 (1988).
- Kurtz, S. K. & Perry, T. T. *J. appl. Phys.* **39**, 3798–3813 (1968).
- Blinov, L. M. et al. *Liq. Cryst.* **2**, 121–130 (1987).

ACKNOWLEDGEMENTS. We thank L. A. Beresnev for supplying conventional ferroelectrics, L. Bosio for X-ray measurements, D. Subachyus for pyroelectric measurements, V. Lemoine for his contribution to nonlinear optical experiments and J. Prost for discussions.

Lead isotope evidence for young trace element enrichment in the oceanic upper mantle

Alex N. Halliday*, Gareth R. Davies*, Der-Chuen Lee*, Simone Tommasini*, Cassi R. Paslick*, J. Godfrey Fitton† & Dodie E. James†

* Department of Geological Sciences, University of Michigan, Ann Arbor, Michigan 48109-1063, USA

† Department of Geology and Geophysics, University of Edinburgh, Edinburgh EH9 3JW, UK

ISOTOPIC heterogeneity in ocean island basalts has generally been ascribed to processes related to the long-term cycling of mantle material^{1–6}. A recent study of Cameroon line lavas reported higher $^{208}\text{Pb}/^{204}\text{Pb}$ and $^{206}\text{Pb}/^{204}\text{Pb}$ ratios towards the continent/ocean boundary (c.o.b.), but no corresponding increase in $^{207}\text{Pb}/^{204}\text{Pb}$, indicating large *in situ* fractionations of uranium and thorium relative to lead in the upper mantle $\sim 10^8$ years ago⁷. Here we present neodymium, strontium and lead isotope data for a variety of central Atlantic islands, and show that similar offsets in lead isotope ratios are found in lavas from the islands of Madeira and Trinidad. Like the Cameroon line c.o.b. lavas, these lavas are characterized by high U/Pb and Ce/Pb, low K/U and are located in areas of old oceanic lithosphere⁸. But in contrast to the Cameroon line⁷, the Madeira lavas are derived from an enriched MORB-type source with low $^{207}\text{Pb}/^{204}\text{Pb}$ and $^{87}\text{Sr}/^{86}\text{Sr}$ and high $^{143}\text{Nd}/^{144}\text{Nd}$. The lead isotope data can be explained if the U/Pb ratios in the sources are comparable to those observed for the lavas and the U/Pb fractionation occurred at the time of formation of the local oceanic lithosphere. Although we do not have a satisfactory explanation for the U/Pb fractionation, it must have occurred at shallow depths in the mantle, near a spreading ridge; and the resulting enriched source regions have since remained fixed relative to the migrating lithosphere.

Isotopic compositions of continental intraplate mantle-derived magmas commonly vary as a function of lithospheric age^{9,10}. Oceanic lithosphere, however, is young and effectively of uniform age relative to continental lithosphere. As a consequence any isotopic effects generated by *in situ* decay are

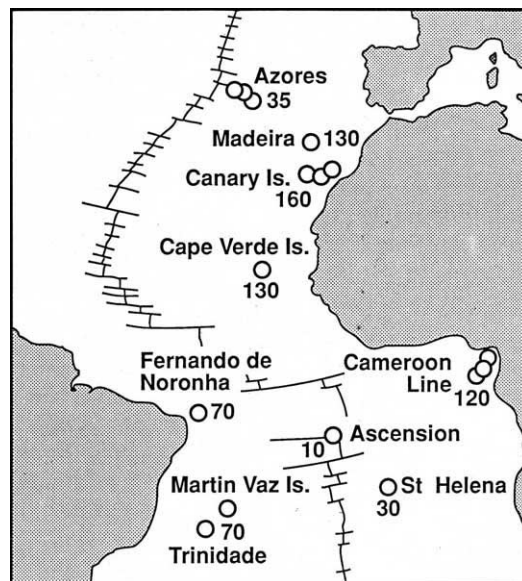


FIG. 1 Map of the central Atlantic Ocean showing locations of OIB volcanic centres. Numbers indicate the age (in Myr) of the local lithosphere⁸.

TABLE 1 isotope data for Central Atlantic OIBs

Sample number	Rock type	MgO (wt%)	$\frac{^{87}\text{Rb}}{^{86}\text{Sr}}$	$\frac{^{147}\text{Sm}}{^{144}\text{Nd}}$	$\frac{^{238}\text{U}}{^{204}\text{Pb}}$	$\frac{^{87}\text{Sr}}{^{86}\text{Sr}}$	$\frac{^{143}\text{Nd}}{^{144}\text{Nd}}$	$\frac{^{206}\text{Pb}}{^{204}\text{Pb}}$	$\frac{^{207}\text{Pb}}{^{204}\text{Pb}}$	$\frac{^{208}\text{Pb}}{^{204}\text{Pb}}$
Madeira										
MD4	alkali basalt	6.5	0.0868	0.1178	40.8	0.702895 ± 13	0.513045 ± 9	19.060	15.440	38.578
MD10	basanite	7.5	0.0780	0.1196	44.8	0.702888 ± 11	0.513063 ± 9	19.208	15.545	38.867
MD32	basanite	12.8	0.0839	0.1233	43.5	0.703032 ± 11	0.513053 ± 8	19.047	15.491	38.683
MD40	alkali basalt	10.7	0.0814	0.1311	35.1	0.702782 ± 11	0.513109 ± 9	19.049	15.490	38.567
MD64	basanite	8.1	0.0838	0.1193	43.0	0.702802 ± 13	0.513113 ± 7	18.949	15.458	38.447
Ascension										
AS1	transitional basalt	5.1	0.1606	0.1344	37.4	0.702847 ± 9	0.513000 ± 8	19.473	15.672	39.190
AS3	transitional basalt	5.3	0.1194	0.1398	22.4	0.702792 ± 9	0.513033 ± 6	19.259	15.619	38.864
AS5	transitional basalt	4.2	0.1737	0.1365	29.4	0.702753 ± 13	0.513058 ± 6	19.431	15.646	39.081
AS10	transitional basalt	4.6	0.1165	0.1323	22.4	0.702913 ± 13	0.513003 ± 5	19.658	15.660	39.311
Trinidad										
TD3	olivine nephelinite	13.4	0.1635	0.1186	26.7	0.703611 ± 11	0.512762 ± 8	19.047	15.564	38.816
TD4	nephelinite	5.6	0.0691	0.0791	52.0	0.703804 ± 13	0.512772 ± 8	19.143	15.535	39.023
TD5	nephelinite	6.6	0.0261	0.1440	44.5	0.703837 ± 9	0.512799 ± 9	19.152	15.523	39.020
Azores-Fayal										
AZFY1	alkali basalt	6.6	0.0732	0.1261	23.2	0.703931 ± 11	0.512863 ± 6	19.153	15.634	38.937
AZFY2	alkali basalt	10.2	0.1329	0.1333	21.1	0.703762 ± 14	0.512930 ± 9	19.058	15.595	38.717
AZFY3	alkali basalt	8.6	0.2059	0.1254	9.62	0.703886 ± 11	0.512878 ± 8	18.637	15.606	38.350
Azores-Pico										
AZP5	transitional basalt	16.2	0.0227	0.1387	20.0	0.703696 ± 14	0.512911 ± 7	19.432	15.610	38.938
AZP6	alkali basalt	6.2	0.1064	0.1249	28.9	0.703443 ± 13	0.512946 ± 6	19.464	15.604	38.936
AZP7	alkali basalt	16.1	0.1146	0.1372	16.6	0.703748 ± 14	0.512885 ± 6	18.755	15.552	38.251
AZP8	alkali basalt	13.2	0.1238	0.1299	19.6	0.703943 ± 14	0.512869 ± 6	19.295	15.676	39.082
Azores-Flores										
AZF1	alkali basalt	4.8	0.1738	0.1107	29.3	0.703380 ± 12	0.512969 ± 7	19.282	15.582	38.990
AZF2	alkali basalt	6.1	0.1806	0.1092	25.8	0.703360 ± 13	0.512934 ± 8	19.289	15.565	38.836
AZF3	transitional basalt	10.6	0.0697	0.1213	26.9	0.703364 ± 8	0.512945 ± 8	19.002	15.546	38.626

All isotopic ratios measured at the University of Michigan using a VG Sector multicollector mass spectrometer. $^{147}\text{Sm}/^{144}\text{Nd}$ and $^{238}\text{U}/^{204}\text{Pb}$ measured by isotope dilution at the University of Michigan. MgO and $^{87}\text{Rb}/^{86}\text{Sr}$ determined by X-ray fluorescence at the University of Edinburgh. Uncertainties in isotopic ratios refer to least significant digits and represent $\pm 2\sigma$ mean run precisions. Sr and Nd isotopic compositions are presented normalized to $^{86}\text{Sr}/^{88}\text{Sr} = 0.1194$ and $^{146}\text{Nd}/^{144}\text{Nd} = 0.7219$, respectively, and were measured in multidynamic mode. Pb isotopic ratios are presented normalized to NBS SRM981 using a fractionation correction of 0.101% per a.m.u. and were measured in static mode. The $^{87}\text{Sr}/^{86}\text{Sr}$ of NBS SRM987 was 0.71025 ± 1 (2σ , $N=20$) and the $^{143}\text{Nd}/^{144}\text{Nd}$ of the La Jolla Nd standard was 0.51185 ± 1 (2σ , $N=20$) at the time of these analyses.

likely to be small compared with those inherited from asthenospheric heterogeneities. The central Atlantic is an ideal location in which to study the relationships between the lithosphere and isotopic compositions in basalts because the range in lithospheric age is large⁸, there are no recent subduction zones nearby and the area is well away from the effects of the Dupal anomaly in the Southern Hemisphere^{5,11}. The locations of the main intra-plate volcanic centres in the central Atlantic region are shown in Fig. 1. The age of the adjacent oceanic lithosphere ranges from close to zero (Ascension) to ~160 Myr (Canary Islands)⁸.

New Nd, Sr and Pb isotope data, together with parent/daughter ratios for basaltic lavas, from several of these islands are presented in Table 1, and plotted in Fig. 2, along with previous data for a variety of other ocean island basalts (OIBs) and mid-ocean-ridge basalts (MORBs)^{12,19}. Details of the geology of the islands studied here can be found elsewhere (refs 8, 13–15, 20–24 and refs therein). The least well known are Trinidad and Madeira. Trinidad is a volcanic cone rising 5,500 m above sea level and is dominated by rocks undersaturated in silica, in the form of pyroclastic deposits, domes, dykes and plugs, together with subordinate lava flows that include nephelinite, ankartrite and phonolite, all of which are dated at <4 Myr (ref. 21). Madeira is the main island of an archipelago, and the rocks are dominated by alkaline olivine basalts and their differentiates (chiefly hawaiites and mugearites) that are unusually rich in Na, and have been dated at >3 Myr (refs 22, 23).

The Nd and Sr isotopic compositions for Madeira, the Cameroon line and Trinidad, situated in areas of old lithosphere, are not significantly different from those for islands closer to the Mid-Atlantic Ridge such as the Azores, Ascension and St Helena (Fig. 2a). For example, both Ascension and

Madeira are characterized by very radiogenic Nd and unradiogenic Sr plotting within the MORB field, and the Cameroon line c.o.b. samples show similarities to HIMU or Lo-Nd islands such as St Helena. But although Madeira, Trinidad and the Cameroon line c.o.b. show correlated variations in $^{208}\text{Pb}/^{204}\text{Pb}$ and $^{206}\text{Pb}/^{204}\text{Pb}$ and plot within the general MORB–OIB array (Fig. 2b), there is no correlation with $^{207}\text{Pb}/^{204}\text{Pb}$. Instead, these volcanic centres define fields that are displaced to the right of the array for other OIB and MORB in Fig. 2c. This is evidence of both extreme and young source enrichments in U and Th relative to Pb (ref. 7). The shorter half-life of ^{235}U results in it having a much lower abundance at present relative to ^{238}U . Because U and Th are believed to have similar distribution coefficients during mantle melting, a large fractionation in U/Pb is associated with change of similar magnitude in Th/Pb. The time-integrated effect of these fractionations in the recent past (say 100 Myr ago) will be correlated changes in $^{206}\text{Pb}/^{204}\text{Pb}$ and $^{208}\text{Pb}/^{204}\text{Pb}$, at relatively constant $^{207}\text{Pb}/^{204}\text{Pb}$. Because U and Th are more incompatible than Pb during mantle melting^{6,7,16,25–27}, limits can be placed on the time needed to generate the offsets shown in Fig. 2c, given an upper limit for the source U/Pb ratio corresponding to the measured values for the lavas (Table 1). The lavas of these three regions are characterized by unusually high measured U/Pb (Fig. 3a). It is possible to achieve the largest offsets in Pb isotopic composition at Madeira with U/Pb as high as measured in the lavas in ~130 Myr, the age of the lithosphere (Fig. 2c). It is difficult, however, to explain these offsets with much lower U/Pb in the source and longer times, as this enhances the $^{207}\text{Pb}/^{204}\text{Pb}$ as well as the $^{206}\text{Pb}/^{204}\text{Pb}$ ratios. A similar case was previously made for the source regions of the Cameroon line c.o.b.⁷. We therefore interpret the isotope

data for these two regions as indicating that the high U/Pb ratios of the lavas are also features of the source, with the age of U/Pb fractionation being comparable to the age of the lithosphere. The similarity between modelled and measured U/Pb ratios indicates that there was only minor fractionation of U/Pb during the partial melting that produced the lavas, consistent

with the very low mineral-melt partition coefficients that have been measured for U and Pb (refs 25, 26). The Trinidad data display less pronounced offsets despite having high U/Pb, consistent with the younger age of the lithosphere at this location (Fig. 1). But the measured U/Pb ratios are variable and not correlated with $^{206}\text{Pb}/^{204}\text{Pb}$, implying some additional recent U/Pb fractionation, presumably associated with magma generation and differentiation. The Pb isotope data for other islands located in areas of old lithosphere, Fernando de Noronha, Cape Verdes and the oceanic Cameroon line (Fig. 1), tend to be only slightly high in $^{206}\text{Pb}/^{204}\text{Pb}$ for a given $^{207}\text{Pb}/^{204}\text{Pb}$ (Fig. 2c), as expected given the lower average U/Pb ratios of lavas from these islands (Fig. 3a).

The respective portions of precursor mantle for Madeira and the Cameroon line c.o.b. before their enrichment in trace elements were radically different from each other. The Cameroon line basaltic magmas have OIB-like isotopic compositions and are thought to be derived by melting of a fossil plume, the head of which became fractionated in U/Pb about 125 Myr ago^{7,14}. In contrast, the Madeira basalts have $^{87}\text{Sr}/^{86}\text{Sr}$, $^{143}\text{Nd}/^{144}\text{Nd}$ and $^{207}\text{Pb}/^{204}\text{Pb}$ ratios that are similar to modern MORB (Fig. 2a), and among the most extreme yet measured for OIB. Therefore, the most likely origin for the Madeira basalts is in upper-mantle MORB source material that was enriched $\sim 10^8$ yr ago. Clearly the trace-element enrichment process responsible for the offsets in Pb isotopic composition in Fig. 2c is largely independent of the exact precursor material and the longer-term heterogeneity of the mantle¹⁻⁶.

The incompatible element ratio K/U is considered uniform in MORB, a feature that is thought to reflect similar bulk partition coefficients for K and U such that there is negligible fractionation during melting²⁸. In contrast, U is generally considered significantly more incompatible than Pb during melting^{7,16,26,27}. Therefore, if the degree of melting is such that U/Pb ratios of OIB reflect their source compositions, the K/U ratios must also do so, unless there are phases in the OIB source that selectively fractionate K from U. A similar argument can be made for Ce/Pb ratios²⁹. Lavas from Ascension and the Azores, close to the Mid-Atlantic Ridge, have K/U ratios that are similar to average MORB (Fig. 3b), whereas St Helena and centres from areas of old oceanic lithosphere have considerably lower ratios. The centres in oldest lithosphere (>100 Myr) with higher average U/Pb also have lower average K/U, suggesting that the process that fractionates the former may also be responsible for fractionation of the latter. Overall the average K/U and U/Pb ratios define a broad hyperbolic trend that may reflect mixing, but is more likely to indicate a process that fractionates both K and Pb relative to U (Fig. 3b). The centres with high U/Pb also have higher average Ce/Pb than most basalts²⁹ (Fig. 3c). If the U/Pb ratios are representative of their sources and were fractionated at the time of formation of the oceanic lithosphere, this must also be true of the Ce/Pb ratios.

The average Nd isotopic compositions of central Atlantic OIB correlate with their respective average Sm/Nd (Fig. 3d), although the average Madeira composition lies away from the trend defined by the other islands. Unlike combined Pb isotope data, Nd isotopic compositions cannot uniquely define both parent/daughter ratios and time. Assuming that average measured Sm/Nd ratios are lower than those of their source regions, the trend for most of the islands suggests that the age of the Nd isotopic heterogeneity is less than ~ 1 Gyr. To generate the range in Nd isotopic composition over the 130 Myr modelled for the Pb isotope data, however, requires excessively large variation in source Sm/Nd and does not explain the collinearity for islands with different lithospheric age. Therefore the Nd isotope data reflect long-term heterogeneity, and provide evidence that the process that fractionated U/Pb resulted in little fractionation of Sm/Nd. Similarly the melting that produced most of the erupted lavas resulted in a limited and/or systematic fractionation of Sm/Nd. This may be the result of

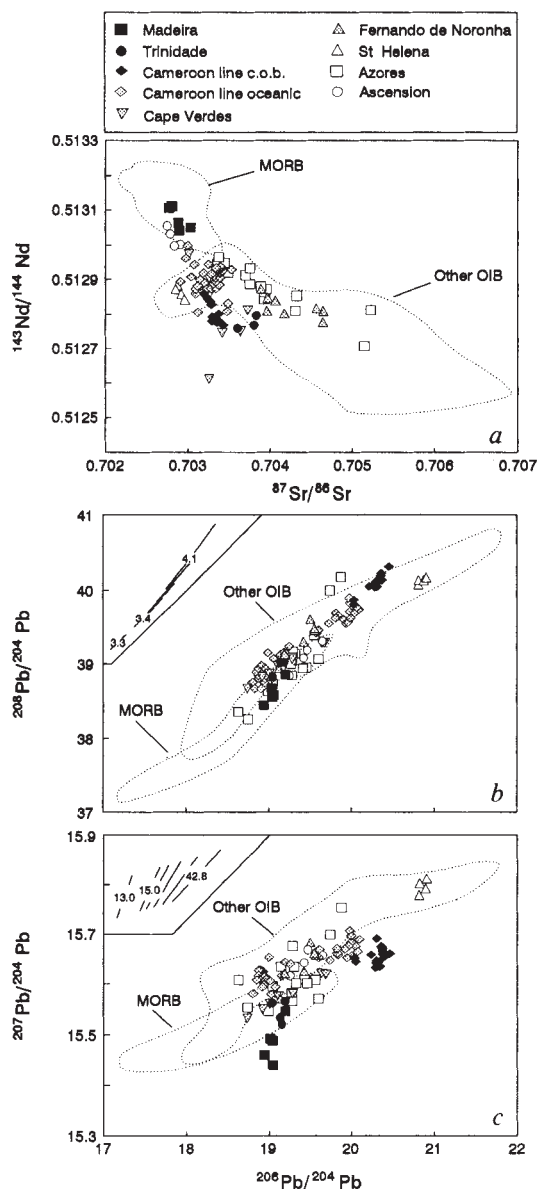
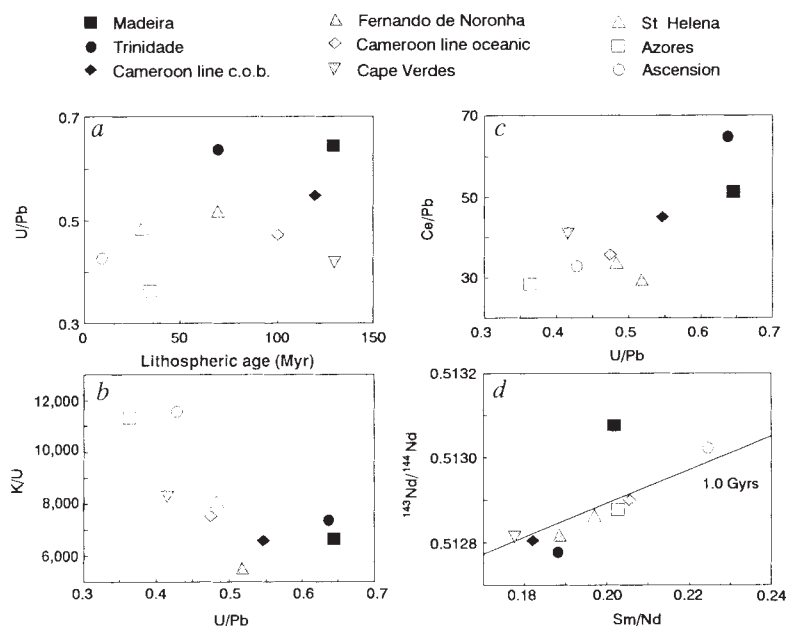


FIG. 2 Plots of *a*, Sr and Nd, and *b*, *c*, Pb isotopic compositions of lavas from the central Atlantic islands plus data for other OIB and MORB from the literature (Table 1 and refs. 7, 12–19). Where possible, we have plotted only samples for which U, Pb, Nd, and Sm data from isotope dilution, and major- and trace-element data from X-ray fluorescence were available. Comparisons with published data for central Atlantic islands are restricted to samples of common rock types, mainly transitional basalts, alkali basalts and basanites with a few nephelinites. Inserts to *b* and *c* show the changes expected in the Pb isotopic composition of a model array of data over a period of 130 Myr with μ ($^{238}\text{U}/^{204}\text{Pb}$) values comparable to those found in the erupted basalts. The array to the far left mimics the lead isotopic compositions at 130 Myr. The series of arrays to the right are present-day arrays produced by variable increases in μ and κ ($^{232}\text{Th}/^{238}\text{U}$) at 130 Myr. It is assumed that the μ and κ ratios in the precursor source, that is before enrichment, were variable in proportion to their respective Pb isotope ratios. The μ and κ produced in the source during enrichment at 130 Myr vary in proportion to the precursor source ratios, so the resultant Pb isotope arrays have different slopes. The average μ and κ are shown for each of the two most extreme arrays.

FIG. 3 Average U/Pb for each central Atlantic island as a function of lithospheric age; *b*, average K/U ratio against U/Pb ratio; *c*, average Ce/Pb ratio against U/Pb ratio and *d*, Nd isotope composition against Sm/Nd ratio. Data are from refs 7, 12–16, 20, 24 from Table 1 and our unpublished work.



low-pressure melting. The deviation of the average Madeira composition from the correlation for the other islands supports the view that the origins of the Madeira magmas are distinct in being derived by small degrees of partial melting at high pressure in the garnet stability field, as garnet would cause fractionation of the rare earth elements²³.

Several processes may be considered potentially responsible for the fractionations in U/Pb, Ce/Pb and K/U at the time of lithospheric growth. The melt regimes under ridges and hotspots are usually modelled as regions that are zoned laterally and vertically with respect to degree of partial melting^{4,27,30–32}. Plume heads may be re-enriched by percolation of small-degree partial melts, such as nephelinitic magmas^{7,30}. But there is little evidence that the U/Pb ratios of the source regions of the lavas studied here were fractionated simply by enriched small-degree partial melts. The inverse correlation of measured Sm/Nd and U/Pb for individual samples in the Cameroon line⁷, and Trinidad (Table 1), implies that the fractionation of these trace element ratios is linked, but the average measured U/Pb and Sm/Nd ratios shown in Fig. 3 are not correlated, although both are source-related. Furthermore, small-degree partial melting in the presence of only the main upper-mantle mineral phases (assuming plagioclase is not involved) should not produce a positive correlation between U/Pb and Ce/Pb (Fig. 3c) on the basis of published mineral-melt distribution coefficients^{25–27}. Minor buffering phases such as phlogopite and carbonates, which may play a role in low-degree partial melting of the mantle^{6,30}, have been invoked as an explanation for high U/Pb in areas of old continental lithosphere¹⁰, but there is no obvious phase that will buffer K and Pb to produce the effect observed in Fig. 3b, without also affecting Ba or Ti; Ba/Ce ratios of central Atlantic OIB are in fact relatively uniform. Low-viscosity carbonatitic liquids may separate from immiscible nephelinitic liquids at small degrees of melting with concomitant fractionation of incompatible elements. This might explain the extreme U/Pb and Sm/Nd fractionations found in Trinidad and Cameroon line nephelinites. Nevertheless, although carbonatitic liquids may transport incompatible elements efficiently through the mantle³², calculated compositions using experimental partition coefficients³³ cannot reproduce the fractionations shown in Fig. 3. A more speculative explanation is that there is a low-density, low-viscosity, subsolidus fluid phase associated with nephelinite magma generation but at greater depth below the melt zone, analogous to that suggested for Hawaii³⁰ into which, in the absence of phlogopite, K and Pb may preferentially partition³⁴

relative to U and the light rare earth elements. This fluid phase may migrate and be consumed in melting reactions elsewhere, or rise as a low-viscosity fluid³⁵, causing metasomatism in refractory lithosphere away from the melt zone. The low concentrations of volatiles in MORB glasses³⁶ argue against this explanation.

Whatever the exact mechanism for U/Pb fractionation, it seems to occur at shallow depths in the mantle at about the time of formation of the local oceanic lithosphere. The mantle currently being sampled at a distal ocean island must have undergone fractionation of U/Pb when it was located close to a cooling ridge (in the case of Madeira) or a cooling plume head (in the case of the Cameroon line). The sources of the Cameroon line, Trinidad and Madeira magmas have clearly resided in the upper mantle since the oceanic lithosphere was formed, and have not been displaced laterally by asthenospheric flow relative to the lithosphere since that time⁷. These portions of the mantle may now be melted preferentially in relatively undiluted form, away from the present ridge, by rising plumes³⁷ or hot zones¹⁴ because they are enriched in low-temperature melting fractions incorporated at the time of U/Pb fractionation, and because of a lower geothermal gradient in areas of old thick lithosphere where the peridotite solidus is at greater depth. Elsewhere, such components are diluted by the larger volumes of magma generated at ridges and by rising plumes capable of ascending to shallower depths where the lithosphere is thin. The trace-element fractionations cannot, however, be explained simply as the product of small degrees of partial melting in areas of thicker lithosphere. Rather, storage of fractionated components for tens of millions of years is required. Such a model is endorsed by ³He/⁴He ratios lower than in MORB for the Cameroon line³⁸, suggesting a history of U enrichment. Finally, it should be noted that high U/Pb does not survive for long in the convecting mantle; otherwise the magnitude of Pb isotopic heterogeneity in basalts^{1–6} would be greatly increased. □

Received 30 March; accepted 20 August 1992.

- Hofmann, A. W. & White, W. M. *Earth planet. Sci. Lett.* **57**, 421–436 (1982).
- McKenzie, D. & O'Nions, R. K. *Nature* **301**, 229–231 (1983).
- Allègre, C. J., Hamelin, B., Provost, A. & Dupré, B. *Earth planet. Sci. Lett.* **81**, 319–337 (1986).
- Galer, S. J. G. & O'Nions, R. K. *Chem. Geol.* **56**, 45–61 (1986).
- Zindler, A. & Hart, S. A. *Rev. Earth planet. Sci.* **14**, 493–571 (1986).
- Galer, S. J. G. & O'Nions, R. K. *Nature* **316**, 776–782 (1985).
- Halliday, A. N. *et al. Nature* **347**, 523–528 (1990).
- Emery, K. O. & Uchupi, E. *The Geology of the Atlantic Ocean* (Springer, New York, 1984).

9. Vollmer, R. & Norry, M. J. *Nature* **301**, 141–143 (1983).
10. Hawkesworth, C. J., Kempton, P. D., Rogers, N. W., Ellam, R. M. & van Calsteren, P. W. *Earth planet. Sci. Lett.* **96**, 256–268 (1990).
11. Hart, S. R. *Nature* **309**, 753–757 (1984).
12. Sun, S.-S. & McDonough, W. F. in *Magmaism in the Ocean Basins* (ed. Saunders, A. D. & Norry, M. J.) 313–345 (Geol. Soc. Spec. Publ. No. 42, Blackwell Scientific, Oxford, 1989).
13. Davies, G. R., Norry, M. J., Gerlach, D. C. & Cliff, R. A. in *Magmaism in the Ocean Basins* (ed. Saunders, A. D. & Norry, M. J.) 231–255 (Geol. Soc. Spec. Publ. No. 42, Blackwell Scientific, Oxford, 1989).
14. Halliday, A. N., Dickinson, A. P., Fialick, A. E. & Fitton, J. G. *J. Petrol.* **29**, 181–211 (1988).
15. Gerlach, D. C., Cliff, R. A., Davies, G. R., Norry, M. & Hodgson, N. *Geochim. cosmochim. Acta* **52**, 2979–2992 (1988).
16. Newsom, H. E., White, W. M., Jochum, K. P. & Hofmann, A. W. *Earth planet. Sci. Lett.* **80**, 299–313 (1986).
17. Ito, E., White, W. M. & Gopel, C. *Chem. Geol.* **62**, 157–176 (1987).
18. Palacz, Z. A. & Saunders, A. D. *Earth planet. Sci. Lett.* **79**, 270–280 (1986).
19. White, W. M. & Hofmann, A. W. *Nature* **196**, 821–825 (1982).
20. Gerlach, D. C., Stormer, J. C. Jr & Mueller, P. A. *Earth planet. Sci. Lett.* **85**, 129–144 (1987).
21. Mitchell-Thomé, R. G. *Geology of the South Atlantic Islands* (Gerbrüder Borntraeger, Berlin, 1970).
22. Mitchell-Thomé, R. G. *Geology of the Middle Atlantic Islands* (Gerbrüder Borntraeger, Berlin 1976).
23. Hughes, D. J. & Brown, G. C. *Contrib. Mineral. Petrol.* **37**, 91–109 (1972).
24. Fitton, J. G. & Dunlop, H. M. *Earth planet. Sci. Lett.* **72**, 23–38 (1985).
25. Seitz, M. G. *Yb Carnegie Instn Wash.* **72**, 551–553 (1973).
26. Watson, E. B., Ben Othman, D., Luck, J.-M. & Hoffmann, A. W. *Chem. Geol.* **62**, 191–208 (1987).
27. McKenzie, D. & O'Nions, R. K. *J. Petrol.* **32**, 1021–1091 (1991).
28. Jochum, K. P., Hofmann, A. W., Ito, E., Seufert, H. M. & White, W. M. *Nature* **306**, 431–436 (1984).
29. Hofmann, A. W., Jochum, K. P., Seufert, H. M. & White, W. M. *Earth planet. Sci. Lett.* **79**, 33–45 (1986).
30. Wyllie, P. J. *J. Geophys. Res.* **93**, 4171–4181 (1988).
31. McKenzie, D. *Earth planet. Sci. Lett.* **74**, 81–91 (1985).
32. McKenzie, D. *Earth planet. Sci. Lett.* **72**, 149–157 (1985).
33. Fielding, K. D., Fitton, J. G., Ford, C. E. & Hinton, R. W. *Terra Abstr.* **3**, 419 (1991).
34. Schneider, M. E. & Egger, D. H. *Geochim. Cosmochim. Acta* **50**, 711–724 (1986).
35. Watson, E. B., Brennan, J. M. & Baker, D. R. in *Continental Mantle* (ed. Menzies, M.) 111–124 (Clarendon, Oxford, 1990).
36. Michael, P. J. *Geochim. cosmochim. Acta* **52**, 555–566 (1988).
37. Richards, M. A. & Griffiths, R. W. *Nature* **342**, 900–902 (1989).
38. Sano, Y., Wakita, H., Ohsumi, T. & Kusakabe, M. *Geophys. Res. Lett.* **14**, 1039–1041 (1987).

ACKNOWLEDGEMENTS. We thank K. O'Nions, R. Lange and Y. Zhang for discussions, colleagues in the Radiogenic Isotope Geochemistry Laboratory for criticism of the text, and W. White for a review. The research was supported by NSF grants to A.N.H.

A sharp and flat section of the core–mantle boundary

John E. Vidale & Harley M. Benz

United States Geological Survey, Branch of Seismology,
345 Middlefield Road MS 977, Menlo Park, California 94025, USA

THE transition zone between the Earth's core and mantle plays an important role as a boundary layer for mantle and core convection¹. This zone conducts a large amount of heat from the core to the mantle, and contains at least one thermal boundary layer^{2,3}; the proximity of reactive silicates and molten iron leads to the possibility of zones of intermediate composition⁴. Here we investigate one region of the core–mantle boundary using seismic waves that are converted from shear to compressional waves by reflection at the boundary. The use of this phase (known as ScP), the large number of receiving stations, and the large aperture of our array all provide higher resolution than has previously been possible^{5–7}. For the 350-km-long section of the core–mantle boundary under the northeast Pacific sampled by the reflections, the local boundary topography has an amplitude of less than 500 m, no sharp radial gradients exist in the 400 km above the boundary, and the mantle-to-core transition occurs over less than 1 km. The simplicity of the structure near and above the core–mantle boundary argues against chemical heterogeneity at the base of the mantle in this location.

The regional seismic networks of the United States and Canada routinely record local and distant earthquakes, mainly in order to monitor active faults within the networks. Together, these networks include more than 1,500 stations and span a wide aperture. This set of receivers is ideal for observing short-period seismic waves from distant earthquakes across larger distances than previously possible⁸.

The phase ScP travels from the earthquake to the core–mantle boundary (CMB) as an S wave, then converts to a P wave for the remaining path to the receiver (Fig. 1a). ScP appears in a

particularly quiet interval on seismograms for an earthquake that occurred on 13 March 1992, 180 km beneath the Andreanof Islands (magnitude m_b 6.1, 52.8° N, 178.8° W). We analyse seismograms from the University of Washington Seismic Network (125 stations), the Southern California Seismic Network (216 stations) and the Northern California Seismic Network (405 stations). The high signal-to-noise ratios of these seismograms will reveal the degree of lateral variation in structure near the CMB, the presence or absence of structure above the CMB and the sharpness of the CMB. These data are only sensitive to lateral variations in the CMB with wavelengths of 50 to 200 km, these limits arising from the limitations in resolution of 1-Hz seismic waves and the 15° aperture of the array, respectively. The data are sensitive to radial variations with wavelengths up to 10 km, beyond which insufficient short-period energy is reflected.

A smaller (m_b 5.3) event on 27 March 1992 (52.8° N, 174.0° W, 190 km depth) produced a particularly sharp set of arrivals, but was only available on the Northern California Seismic Network. The simplicity of the earthquake allows a direct measure of the simplicity of the CMB.

ScP is prominent in the distance range from 30° to 70° because the conversion coefficient at the CMB is large. The main source of short-period noise in this time interval is the coda of the P wave, which attenuates rapidly with time after the P wave passes. We study the CMB with the ScP phase rather than the more traditional choice of PcP^{5–7} (Fig. 1) because ScP arrives 4 minutes later with a more favourable signal-to-noise ratio.

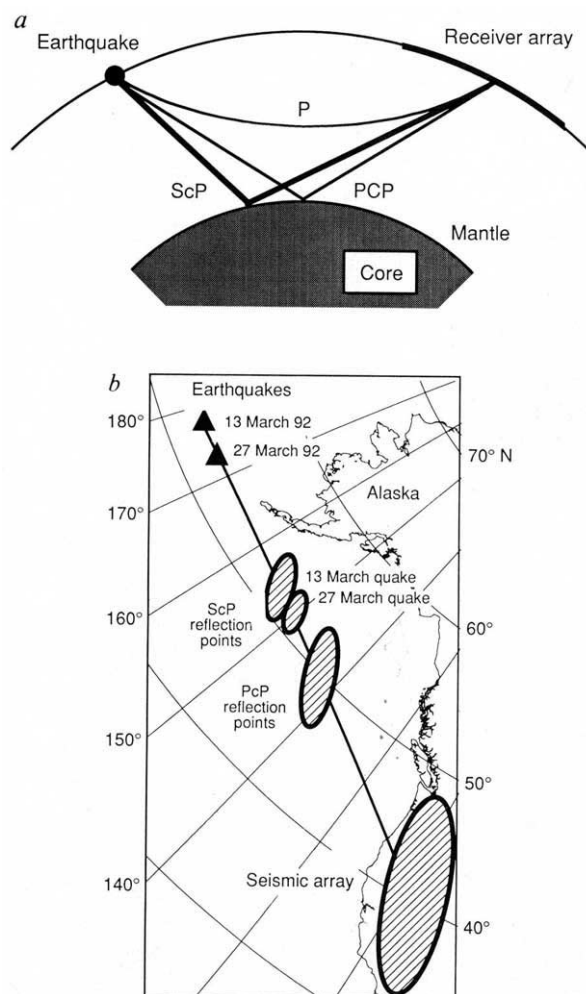


FIG. 1 Schematic ray paths (a) and bounce points at the core–mantle boundary (b) for the seismic waves analysed here.

We selected the ScP arrivals at the 362 out of 746 stations with the highest signal-to-noise ratio. These phases were aligned, then the seismograms were summed for 1° distance intervals (Fig. 2a), revealing nearly identical waveforms across the array. The ScP arrival lasts ~ 5 s, typical of the source duration for magnitude 6.1 earthquake. The complexity caused by the source duration is then removed by first estimating the source time function to be the sum of all the aligned traces, then deconvolving this sum from each trace. Deconvolved traces summed in 1° bins are shown in Fig. 2b. Again, the waveforms are nearly identical. The smaller earthquake, 400 km closer to the array and at a slightly different azimuth, is a more impulsive source (Fig. 2c). It samples a 130-km patch of the CMB adjacent to the 250-km patch sampled by the first earthquake. The waveform of the ScP phase does not vary across the array.

For the smaller earthquake, we aligned and summed the P waves, aligned and summed the ScP waves, and summed the PcP waves aligned according to ScP with the expected decrease in apparent velocity across the array (Fig. 3a). The ability to stack more than 100 seismograms eliminates the near-receiver crustal reverberations that have complicated previous comparisons of P with PcP^{5,6}. Only the P and ScP arrivals contained enough energy in the interval 0.5 to 10 s to allow removal of the instrument response (Fig. 3b). The similarity of the P, PcP and ScP frequency content requires that the shear wave attenuation is low along the path from the earthquake to the CMB, which has also been observed for other paths⁹.

Overall, the ScP, P and PcP waveforms are similar, particularly in the first second, indicating that the CMB is primarily a sharp transition. Our observations do not show differences as large as

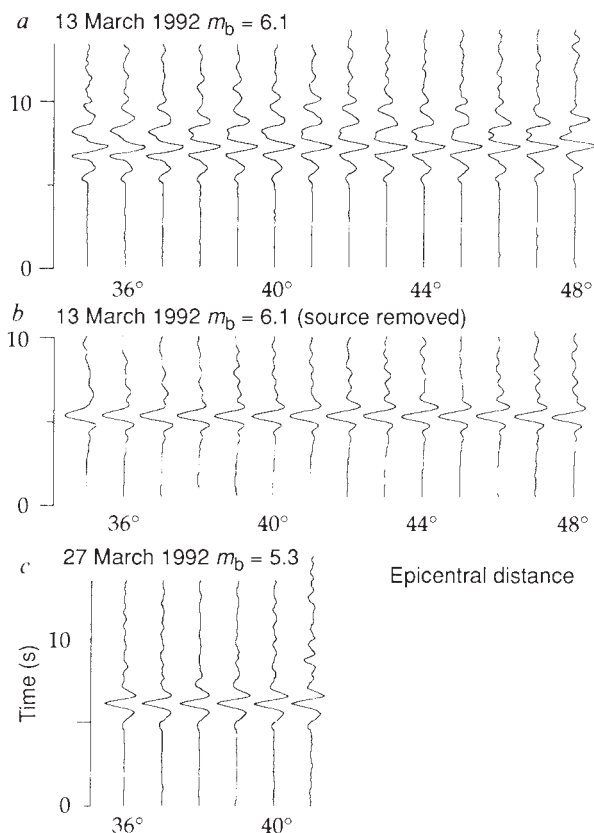


FIG. 2 Seismograms summed over 1° distance intervals are shown after alignment of the ScP phases for *a*, the 13 March 1992 Andreanof earthquake; *b*, deconvolved records of the 13 March 1992 Andreanof earthquake; and *c*, the 27 March 1992 Andreanof earthquake. These traces are from short-period, vertical-component seismograms that have usable signal from 0.5 to 10 s period for these earthquakes. The traces show the similarity of the ScP phases.

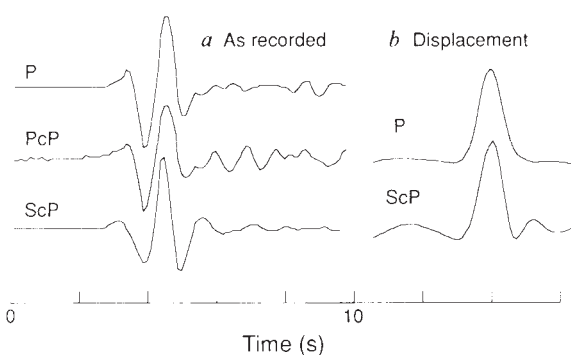


FIG. 3 *a*, Waveforms of the P, PcP and ScP waves of the 27 March 1992 earthquake ($m_b = 5.3$) estimated by the summation of 110–120 clean records from the Northern California Seismic Network. *b*, P and ScP displacement time histories. A 12-s period sine wave was added to cancel the long-period motions not resolved by our limited passband (see Fig. 2 legend).

those that have been interpreted in noisier data as evidence for a laminated CMB⁶.

The similarity of the waveforms in Figs 2 and 3 places tight limits on possible CMB topography. Calculations for a CMB with 500 m of relief^{10,11} show resolvable waveform complexity that varies with epicentral distance and azimuth. Our observations therefore require less than 500 m topography with 50 to 200 km wavelength, although we can not preclude topography with larger or smaller scale. The small variations that are visible in the later half of the ScP waveform in Fig. 2a and *b* may arise from a few hundred metres of CMB topography or velocity variations above the CMB. Source directivity is a less likely explanation as the entire P-wave arrival is similar across the array, and a greater range of take-off angles is spanned by P than by ScP.

The P wave in Fig. 3b is very simple, showing that the earthquake lasted only 1 to 2 s. The ScP wave has one more half cycle than P and PcP in Fig. 3a, and this translates into the small secondary peak on the displacement record. The extra half cycle is visible for all bins in Fig. 2c. This complexity is difficult to attribute to extra attenuation of the S wave, because attenuation would cause the higher frequencies to arrive earlier, not later as observed. It is also difficult to attribute the complexity to lateral heterogeneity near the CMB, as a reflection from a point scatterer at or above the CMB should shift noticeably in time across the array. The difference may be due to a difference between the P- and S-wave radiation from the source.

Structures above the CMB would generate precursors to ScP. We expect structure, if present, to be above the CMB, as little heterogeneity is expected in the liquid core¹². The bounce points of the ScP arrivals shown here are interesting for two reasons. First, this region beneath southern Alaska shows clear evidence for an anomalously fast D'' layer from long-period shear waves¹³. The top of the D'' layer is ~ 250 km above the CMB. We see no evidence of reflections from this height, which may indicate that the transition at the top of D'' is rather gradual. Second, arrivals between P and PcP that are identifiable on individual traces have been interpreted¹⁴ as arising from intermittent layering above the CMB just a few hundred kilometres away from our bounce points, yet we do not see these arrivals on our traces.

The interval before the arrival of ScP from the larger, 13 March earthquake is clean. No coherent energy is visible in the complete seismic section, nor in the numerous sections from subsets of the data that we examined. The noise levels are 5–20% of the ScP amplitude, and in our experience, arrivals as small as one-half of the noise level are visible in such dense record sections. As a more sensitive test for layered reflectors, we aligned the ScP arrivals, then summed the traces. The noise level is near 0.5%. There are no clear short-period reflections from the top nor from within D''.

The sharpness of the CMB may be estimated from the reflection coefficient of ScP as a function of frequency. It is apparent from high-pass filtered traces that 2-Hz energy is reflected from the CMB along our entire profile. And the duration of the 2-Hz arrivals is ~ 5 s, close to the duration of the earthquake. The presence of 2-Hz reflections shows that the CMB has a significant vertical velocity gradient spanning less than 1 km across the entire profile.

We conclude that within our resolution, the CMB is flat, sharp, lacks lateral variation and is not overlain by structures that would reflect 0.5- to 2-Hz energy, at least for a 350-km span beneath the northeast Pacific Ocean.

The lack of topography greater than 500 m with a scale length of 50–200 km directly confirms inferences drawn from PcP amplitudes^{10,11}, PKP precursors¹⁵ and length of day fluctuations¹⁶. The boundary sharpness required along this section of the CMB, with the change from mantle to core occurring across less than 1 km, is more stringent than previous measurements from individual records of P and PcP^{5,6}.

The lack of overlying structure is in contrast to the results of long-period studies. Long-period shear wave studies worldwide indicate an intermittent 2–5% increase in impedance with depth roughly 250 km above the CMB^{17,18}. Near Alaska, at the location of our bounce points, the 2–3% increase in the shear wave velocity has been proposed¹⁹ to occur in a depth range of less than 25 km. The long-period measurements show that the D'' layer is faster than normal mantle^{17,20}. This argues against its being interpreted solely as a thermal boundary layer, as a hot D'' would be slow². Our lack of observed short-period reflections

requires that such a boundary be broader than 5 km. The absence of critical reflections from the top of D'' also argues for a gradual boundary²¹, as mentioned above. A remote possibility is that the boundary is so irregular that short-period energy reflects incoherently, whereas long-period reflections can be observed. Thus, either D'' is absent in the region that we sample, or its upper boundary is unexpectedly diffuse. □

Received 5 June; accepted 28 August 1992.

1. Lay, T. *Eos* **70**, 54–55; 58–59 (1989).
2. Jeanloz, R. & Richter, F. M. *J. geophys. Res.* **84**, 5497–5504 (1979).
3. Williams, O. & Jeanloz, R. *J. geophys. Res.* **95**, 19299–19310 (1990).
4. Knittle, E. & Jeanloz, R. *Geophys. Res. Lett.* **16**, 609–612 (1989).
5. Kanamori, H. *J. geophys. Res.* **72**, 559–572 (1967).
6. Niaz, M. & McLaughlin, K. L. *J. Geodynam.* **8**, 1–16 (1987).
7. Chowdhury, D. K. & Frasier, C. W. *J. geophys. Res.* **78**, 6021–6028 (1973).
8. Vidale, J. E. & Benz, H. M. *Nature* **356**, 678–683 (1992).
9. Burdick, L. J. *Geophys. J. R. astr. Soc.* **80**, 35–55 (1985).
10. Menke, W. *Geophys. Res. Lett.* **13**, 1501–1504 (1986).
11. Kampfmann, W. & Müller, G. *Geophys. Res. Lett.* **16**, 653–656 (1989).
12. Stevenson, D. J. *Geophys. J. R. astr. Soc.* **88**, 311–319 (1987).
13. Young, C. J. & Lay, T. *Rev. Earth planet. Sci.* **15**, 25–46 (1987).
14. Weber, M. & Koernig, M. *Geophys. Res. Lett.* **17**, 1993–1996 (1990).
15. Bataille, K. & Flatté, S. M. *J. geophys. Res.* **93**, 15057–15064 (1990).
16. Gwinn, C. R., Herring, T. A. & Shapiro, I. I. *J. geophys. Res.* **91**, 4755–4767 (1986).
17. Lay, T. & Helmberger, D. V. *J. geophys. Res.* **88**, 8160–8170 (1983).
18. Revenaugh, J. & Jordan, T. H. *J. geophys. Res.* **96**, 19811–19824 (1991).
19. Young, C. J. & Lay, T. *J. geophys. Res.* **95**, 17385–17402 (1990).
20. Tanimoto, T. *Geophys. J. int.* **100**, 327–336 (1990).
21. Schlittenhardt, J. *J. geophys.* **60**, 1–18 (1986).

ACKNOWLEDGEMENTS. We thank S. Grand, T. Lay and Q. Williams for discussions and H. Houston, C. Ammon, S. Kirby, M. Weber and J. Mori for comments. X. Y. Ding estimated effects of attenuation from reference Earth models. J. Mori removed the instrument response for Fig. 3. R. Crosson demonstrated the effectiveness of source deconvolution. R. Benson and R. Somera assisted in recovering the data.

Blade-shaped conodont elements functioned as cutting teeth

Mark A. Purnell* & Peter H. von Bitter

Department of Invertebrate Palaeontology, Royal Ontario Museum and University of Toronto, 100 Queen's Park, Toronto, Ontario M5S 2C6, Canada

CONODONTS were small eel-shaped animals with vertebrate affinities^{1–4}. Known almost exclusively from the small, tooth-like, phosphatic elements of their feeding apparatus, they have one of the finest fossil records of any group of organisms. Until recently the identity of the animal to which conodont elements belonged was one of palaeontology's great mysteries; the function of the elements themselves, particularly blade-shaped elements, remains uncertain (compare refs 3 and 4). But recent insights into the conodont skeletal Bauplan⁵ allow the debate over function to be taken beyond arguments of analogy. Here we present a functional analysis of opposed blade-shaped-element pairs as components of an integrated apparatus. From this we conclude that such elements operated as cutting teeth within a grasping and food-processing apparatus.

Much of the debate over the function of conodont elements^{6–11} took place before the anatomy of the animal^{1,2} and the architecture of the skeletal apparatus⁵ were known. Consequently, these analyses were limited by the lack of developmental and structural context for the conodont elements and 'the entire field of functional morphology [was] essentially closed to conodont workers'¹⁰. More recent studies have considered the elements as components of an integrated feeding structure within an eel-shaped animal and two functional models have been proposed. The anterior elements of the conodont apparatus, the S and M elements (Fig. 1), may have functioned together as a

ciliated sieve structure that sorted small food particles to be bruised or crushed gently by the posterior Pa and Pb elements^{12,13} (Fig. 1). Alternatively, it has been suggested^{1,5} that the anterior elements grasped the food, which was then cut and ground by the Pb and Pa elements (Fig. 1).

Some of the most convincing evidence for Pa element function

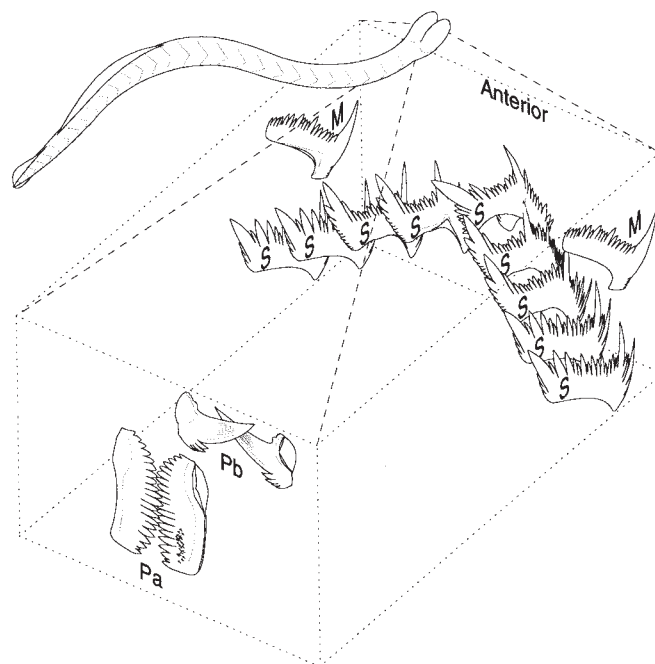


FIG. 1 Feeding apparatus architecture of *Vogelgnathus campbelli* showing the location of the apparatus in the animal. Elements are shown $\times 45$. Relative size is based on bedding plane assemblage specimen 62P-701 (ref. 16). Apparatus architecture is based on ref. 5. S elements are shown more widely spaced than in nature. Animal is shown $\sim \times 1.7$, morphology is diagrammatic, based on refs 1, 2 and 22.

* Present address: Department of Geology, University of Leicester, University Road, Leicester LE1 7RH, UK.

has resulted from physically reconstructing left and right elements into an opposed pair^{13,14} (compare with ref. 15). Commonly, ridges on one element correspond to troughs on the other and the asymmetry in morphology allows intermeshing of elements. Although this methodology has been very illuminating in the analysis of laterally expanded, molar-like elements, no convincing morphological evidence has been found for blade-shaped elements; consequently their function has remained uncertain. They may have been covered in soft tissue, gently rolling small particles of food between their overlapping lateral surfaces^{12,13}, or they may have bitten by occluding⁷, shearing past each other and overlapping⁷, or by hinging backwards in a gate-like biting motion.

Vogelgnathus campbelli (Rexroad) is a well known, widely distributed Lower Carboniferous conodont species that bore paired blade-shaped Pa elements (Fig. 2). Some Pa elements of this species bear small pointed lateral nodes which project upwards at an angle of about 45° (ref. 16; Fig. 2). These lateral nodes are developed only on the anterior left side of the element, irrespective of whether the element is sinistral or dextral (Fig. 2). Although *V. campbelli* Pa elements do not always have lateral nodes, we have studied faunas from Atlantic Canada in which all of them are nodose.

No other species of *Vogelgnathus* bears lateral nodes¹⁷ and there is nothing in the ontogeny or phylogeny of *V. campbelli* to indicate that the nodes were non-adaptive developmental phenomena. It is probable, therefore, that the nodes were functional and that their morphology and arrangement within the apparatus reflect that function. Pa elements in ozarkodinitid

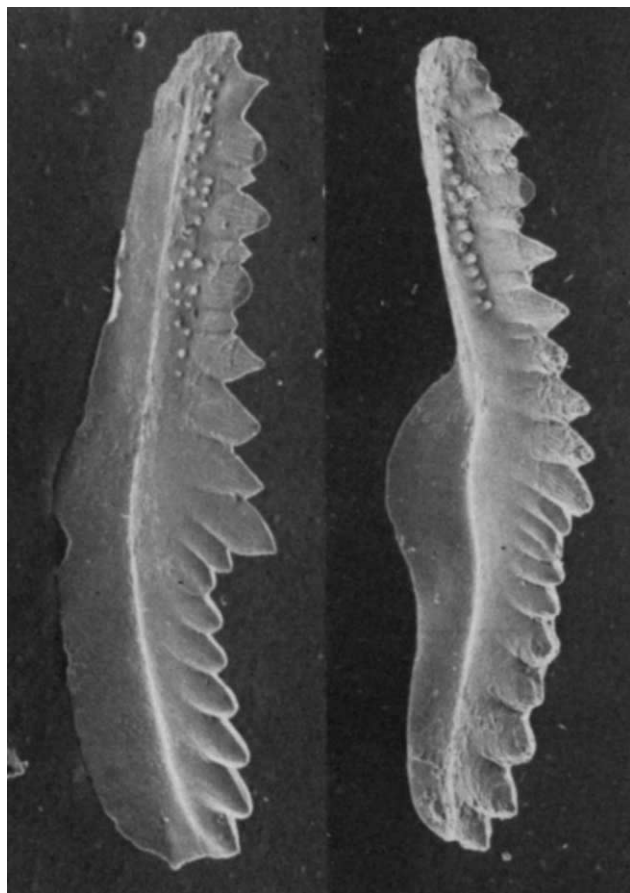


FIG. 2 Scanning electron micrographs of Pa elements of *V. campbelli*, sinistral element on the left, dextral element on the right; oblique upper views showing development of lateral nodes on the anterior left portion of the blade. Sinistral element (specimen number ROM 49379) $\times 300$, dextral element (ROM 49378) $\times 250$.

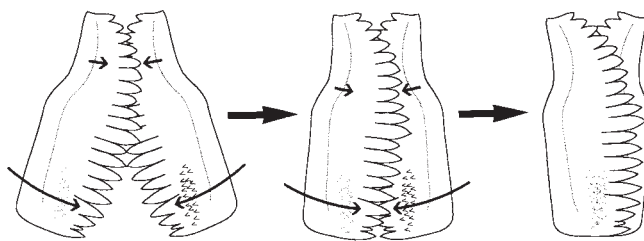


FIG. 3 Operation of *V. campbelli* Pa elements. Note that if elements did not separate further than illustrated, only the lower half of the element (anterior in conventional orientation) could have functioned in cutting. Dotted lateral nodes are seen through the blade of the left element. Elements are shown $\times 70$.

apparatuses functioned as an interactive pair^{5,7,12-14} and we conclude that the lateral nodes also functioned interactively. Thus, the elements were paired with the nodes on the convex side of one element opposed to those on the concave side of the other. The only possible alternative reconstruction places the nodes on the outside of the element pair; a position that precludes interactive function.

Irrespective of the location and arrangement of lateral nodes, if they were permanently covered in soft tissue^{12,13} they would have been cushioned from the physical breakdown of food and could not have functioned. But if the Pa elements functioned as shearing teeth, no such problems arise. The lateral nodes form a structure that closely resembles, and probably functioned as, a rasp that enhanced food processing by moving across the lateral nodes of the opposed element. The convex profile of the elements, the placement of the nodes, and the amount of available space in which they could have moved¹³ suggests that the shearing action of the elements was largely brought about by rotation, with overlap increasing anteriorly (Fig. 3). Neither occluding nor gate-like bites are consistent with the asymmetric placement of nodes on the lateral surfaces of the elements.

Thus, blade-shaped Pa elements in *V. campbelli* sliced and cut food like a pair of serrated scissors. Our analysis indicates that there was a significant functional difference between the inner and outer surfaces of blade-shaped Pa elements (contra ref. 13). A slicing function strongly suggests that the anterior elements of the apparatus (Fig. 1) performed a grasping function as small food particles caught by ciliated sieving elements would not require cutting. These conclusions support recent ontogenetic evidence that growth of the conodont apparatus was incompatible with a filtering function (ref. 18 and M.A.P., manuscript in preparation).

Here we have considered lateral nodes as parts of an interactive pair of elements within the integrated feeding apparatus of an eel-shaped animal. Meaningful interpretation of the nodes is only possible when considered in this context, but this approach also has the benefit of reducing our reliance on assumptions of optimality in function of specific morphological features taken in isolation (one of the main problems in analyses of functional morphology¹⁹⁻²¹). We suggest that a search for asymmetric morphological details in elements of other taxa, coupled with the analytical approach outlined above, can lead to greater understanding of the nature of the conodont apparatus, its evolutionary morphology, and the ecology of individual species □

Received 29 June; accepted 4 September 1992.

1. Briggs, D. E. G., Clarkson, E. N. K. & Aldridge, R. J. *Lethaia* **16**, 1-14 (1983).
2. Aldridge, R. J., Briggs, D. E. G., Clarkson, E. N. K. & Smith, M. P. *Lethaia* **19**, 279-291 (1986).
3. Sansom, I. J., Smith, M. P., Armstrong, H. A. & Smith, M. M. *Science* **256**, 1308-1311 (1992).
4. Briggs, D. E. G. *Science* **256**, 1285-1286 (1992).
5. Aldridge, R. J., Smith, M. P., Norby, R. D. & Briggs, D. E. G. in *Palaeobiology of Conodonts* (ed. Aldridge, R. J.) 63-76 (Horwood, Chichester, 1987).
6. Lindström, M. *Palaeontology* **17**, 729-744 (1974).

7. Jeppsson, L. *Lethaia* **12**, 153–171 (1979).
8. Jeppsson, L. *Lethaia* **13**, 228 (1980).
9. Conway Morris, S. *Lethaia* **13**, 107–108 (1980).
10. Bengtson, S. *Lethaia* **13**, 320 (1980).
11. Bengtson, S. *Lethaia* **16**, 38 (1983).
12. Nicoll, R. S. *BMR J. Aust. Geol. Geophys.* **9**, 133–147 (1985).
13. Nicoll, R. S. in *Palaeobiology of Conodonts* (ed. Aldridge, R. J.) 77–90 (Horwood, Chichester, 1987).
14. Jeppsson, L. *Lethaia* **4**, 101–123 (1971).
15. Weddige, K. *Cour. Forsch. Inst. Senckenberg* **118**, 563–589 (1990).
16. Norby, R. D. & Rexroad, C. B. *Occ. Pap. Indiana geol. Surv.* **50**, 1–14 (1985).
17. Purnell, M. A. & von Bitter, P. H. *J. Paleont.* **66**, 311–332 (1992).
18. Purnell, M. A. *Program Abstr. Can. Paleont. Conf.* **1**, 70 (1991).
19. Gould, S. J. & Lewontin, R. C. *Proc. R. Soc. B* **205**, 581–598 (1979).
20. Signor, P. W. *Neues Jb. geol. paläont. Abh.* **164**, 59–63 (1982).
21. Fisher, D. C. *Paleobiology* **11**, 120–138 (1985).
22. Aldridge, R. J. & Briggs, D. E. G. *Nat. Hist.* **5/89**, 6–10 (1989).

ACKNOWLEDGEMENTS. We thank J. Burke, K. David, and K. Gorra. Supported in part by a grant from the Leverhulme Trust to M.A.P.

Cuckoldry through stored sperm in the sequentially polyandrous spotted sandpiper

Lewis W. Oring*, Robert C. Fleischer†, J. Michael Reed* & Katherine E. Marsden†

* Ecology, Evolution and Conservation Biology Program, University of Nevada, Reno, Nevada 89512, USA

† Genetics Laboratory, National Zoological Park, Smithsonian Institution, Washington DC, Washington 20008, USA

STUDIES of mating systems are often hindered by an inability to determine parentage unequivocally. DNA fingerprinting advanced the field by allowing determination of parentage^{1–3}, especially when alternate mating tactics such as extra-pair fertilizations and parasitic egg-laying are used^{4–6}. A very different, yet essentially unexplored, alternate mating tactic involves fertilization by the sperm of previous mates that is stored for long periods of time. Some birds have the potential to be fertilized by sperm stored in the oviduct for over a month⁷, but studies of long-term sperm storage among wild birds are limited. In the polyandrous spotted sandpiper (*Actitis macularia*), territorial females pair with, defend and lay clutches for several males in rapid succession. Here we report that males pairing early in the season cuckold their females' later mates by means of stored sperm. Thus, early-pairing males not only have greater confidence of paternity, but also increase their reproductive success by fertilizing a proportion of eggs laid for, and incubated by, their females' subsequent mates, in addition to those they incubate themselves.

We studied a population of sex-role reversed spotted sandpipers at Leech Lake, Minnesota, USA. Females arrive about a week before males and fight to establish multipurpose territories to which they attract mates^{8–10}. Males provide most, or all, of the incubation and brood care¹¹. Some females, especially yearlings, are monogamous and help their mates with parental care. But most females, especially those aged 2 years or more, pair with new males, leaving their primary mates to care for the first clutch (classical polyandry)^{12–14}. These females might copulate with, and lay eggs for, up to four different males in as many weeks. Polyandrous females are ideal subjects for testing the likelihood of long-term sperm storage resulting in fertilizations because they copulate repeatedly with each of the males to which they are serially paired.

In most bird species studied, extra-pair copulation (EPC) attempts are concentrated in the female's prelaying period¹⁵. In contrast, EPCs and attempts in spotted sandpipers occur with equal frequency throughout the breeding season (1980–1985, $n = 1,528$ copulations or attempts, $n = 174$ EPCs or attempts)¹⁵. Males face potential loss of reproductive success from EPCs during and immediately before egg laying, as well as those occurring substantially before egg laying. A male also might be

cuckolded by his female's former mates through stored sperm^{16,17}. In at least one species, red-necked phalaropes (*Phalaropus lobatus*), the potential for cuckoldry by means of sperm storage seems to constrain the incidence of polyandry in that males, in choosing mates, discriminate against females that have laid previously¹⁸. Generally, polyandrous birds have a relatively high copulation frequency^{19,20} according to observation of three species of *Phalaropus*, that have only brief pair bonds and lack territorial defence. Data from spotted sandpipers, a species that has longer-lasting pair bonds and territory defence in both sexes, indicate an average copulation frequency similar to several monogamous members of the same subfamily (L.W.O., J.M.R. and S. Maxson, manuscript in preparation). Mate guarding could minimize EPCs during a female's current laying period, but only frequent copulation could minimize the likelihood of eggs being fertilized by the stored sperm of previous mates^{21–22}. Male spotted sandpipers do not alter the amount of time involved in agonistic behaviours before and during laying¹¹, and they guard their mates with constant intensity throughout (L.W.O. *et al.*, manuscript in preparation). Copulation frequency (mean, 0.19 copulations per hour; s.d., 0.16; $n = 21$

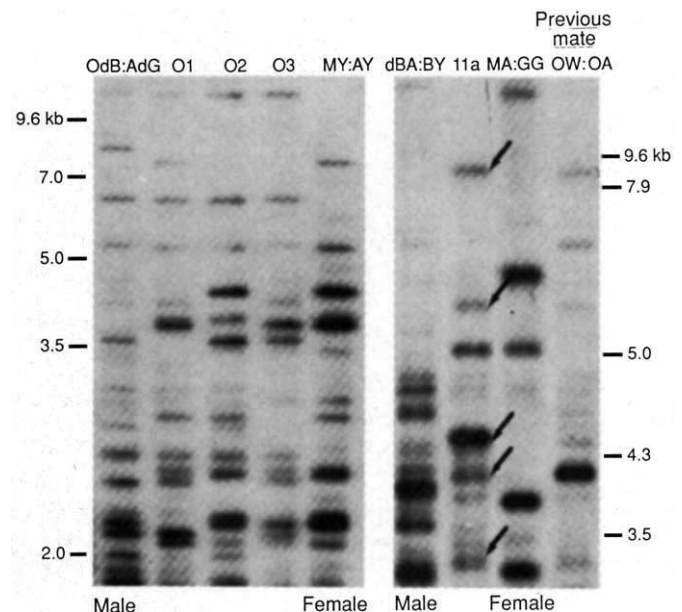


FIG. 1 DNA fingerprint examples (Jeffreys' 33.15 probe) of two spotted sandpiper families. *a*, Fingerprints of female (MY:AY), incubating male (OdB:AdG), and three unexcluded offspring (O1–O3). *b*, Fingerprints showing the female (MA:GG), the incubating male (dBA:BY), one excluded offspring (11a), and the female's previous mate (OW:OA). Note that 11a is excluded from MA:GG and dBA:BY but not from MA:GG and OW:OA (see arrows). METHODS. We analysed blood (200 µl) and embryo samples, stored in 2× standard saline citrate (SSC), obtained from the Leech Lake mainland and from Little Pelican Island, from 25 families (77 young) in 1991 and 9 families (34 young) in 1990. DNA was isolated, dialysed, digested with *Hae*III, and electrophoresed in a 1% agarose gel with 10–20 ng of λ-bacteriophage digested with *Hind*III added to each lane. Digested DNA was depurinated, denatured, and vacuum-transferred to a nylon membrane. The membrane was hybridized to Jeffreys' 33.15 (at 65 °C), and M13 bacteriophage (for a subset of gels, at 60 °C) DNA probes (randomly primed and labelled with [³²P]-dCTP to an activity of 0.5–1.0 × 10⁹ c.p.m. µg⁻¹)²⁷. Filters were washed in 2×SSC:0.1%SDS at the above temperatures and exposed for up to a week. Blots were probed with λ to provide a within-lane size standard to document possible band-shifts. Offspring were excluded if their fingerprint profiles contained more than two fragments not present in parental fingerprint profiles²⁸. The fewest number of offspring bands was 12, but most had 22–35 bands. Offspring listed in Table 1 were excluded from laying female and incubating male (putative parents) by 4–11 extra fragments (mean 7.3) for Jeffreys' 33.15 probe. In the five samples also assessed by M13-generated fingerprints, 3–7 fragments (average of 5.5) excluded the pair from the offspring.

TABLE 1 Sample sizes and exclusions, and mating histories of females that produced paternally excluded embryos or chicks

Total exclusions							
Year	Families	Chicks					
1990	1/9 (11.1%)	1/34 (2.9%)					
1991	6/25 (24.0%)	11/77 (14.3%)					
Total	7/34 (20.6%)	12/111 (10.8%)					
Male exclusions							
Clutch	Female	Male	Clutch complete	Number of eggs	No. of eggs excluded	Unexcluded male	
1990							
1	RA:dBY	dGdG:dGA	30 May	4	1	Unknown	
1991							
1	RA:OB	OR:RA	21 May	4	0	OR:RA	
2		YA:RO	28 May	1	1		
3		YA:RO	10 June	2	0		
4		WA:RY	21 June	1	0		
5		WA:RY	~3 July	4	?		Not collected
1	MA:GG	OW:OA	24 May	4	0	Not collected	
2		dBA:BY	31 May	4	?		Depredated
3		dB:ABY	12 June	2	2		OW:OA
4		dBA:BY	22 June	4	0		
1		OO:dGA	AO:BO	22 May	4		?
2		GO:WA		0	0	Depredated	
3		dGRdG:RA	29 May	2	?		
4		dGRdG:RA	7 June	3	0		
5		AO:BO	13 June	3	0		
6		AO:BO	23 June	4	4		GO:WA
7	MY:AY	AO:BO	~3 July	4	?	Not collected	
1		dBR:AG	28 May	3	0		
2		OR:RA	9 June	4	?		Rotted
3		OdB:AdG	19 June	3	0		
4		AdG:OdG	27 June	3	1		dBR:AG
5	AY:YR	AdG:OdG	~6 July	4	?	Not collected	
1		dGRdG:RA	22 May	4	?		Depredated
2		GO:WA		0	0		
3		MO:OA	3 June	3	0		
4		MO:OA	13 June	3	0		
5		MO:OA	26 June	3	2	dGRdG:RA	
6		MO:OA	~6 July	4	?		Not collected

The number of eggs refers to the number of eggs/embryos collected from the four-egg clutches (when collection occurred). Males mated to females are included even if no clutch was produced.

females, 70 nests; observation time, 1,367.5 pair-hours) varies significantly over the nesting period, peaking the day before the first egg is laid; mated females resist EPC attempts (L.W.O. *et al.*, manuscript in preparation).

In 1990, most samples for DNA analysis were from young, caught and bled after hatching. In 1991, birds were experimentally manipulated to maximize opportunities for extra-pair fertilizations (EPFs) by removing clutches as they were completed. Except for occasions where a clutch was lost immediately, new clutches resulted from pairing with a new mate. Eggs were collected and incubated to develop embryos for DNA analysis. We attempted to synchronize egg removal from nests positioned close to one another so that multiple mates were available to receptive females. Egg removal increased the likelihood of EPFs because more eggs were being laid and because male competition for females was increased. Our clutch removals simulated natural predation when mustelids were resident on the island, a situation that occurred in more than half of the previous 19 years.

We analysed paternity using DNA fingerprinting²³ (Fig. 1). Fingerprints revealed that 1 of 34 offspring (2.9%) in 1990 was excluded from the putative parents. In 1991, 11 of 77 eggs (14.3%) were excluded from their putative parents (Table 1). Analysis of band-sharing indicated that in 11 of 12 cases the incubating males were excluded from paternity (legend, Fig. 1; Table 1). The one remaining offspring probably resulted from an egg laid parasitically. For 10 of the 11 exclusions of the putative male parent, a previous mate of each associated female could not be excluded (Table 1). EPFs were less frequent during

first pairings (3.2% compared with 15.4%; Rice's conditional binomial exact test²⁴, 1-tailed, $P = 0.037$).

Although sperm competition has been studied in domestic birds, little is known about birds in the wild. In domestic birds, sperm from the most recent copulation has the greatest probability of fertilizing eggs⁷. Some fertilizations result from sperm stored for up to 52 days depending on the species⁷. Mated males of some species seem to alter copulation frequency and timing to decrease EPFs^{18,22,25}, but the effectiveness of this behaviour in natural populations is not well known.

Most studies of sperm precedence have looked at fertilization success of two males after a single copulation from each⁷. This is analogous to rapid mate switching in natural settings^{25,26}. In contrast, spotted sandpiper copulations often followed a pattern of repeated copulations by one male, laying of the first clutch, and then repeated copulations by a second male, followed by the laying of a second clutch. Our data indicate that the primary source of EPFs is stored sperm originating from a female's former mates, on the basis of the following evidence. (1) Five former mates, accounting for 10 of 11 EPFs, could not be excluded. Two of five presumed fathers, identified by DNA fingerprinting, disappeared from the island study area and surrounding lake basin 16 and 18 days before the laying of the excluded eggs and were not seen again that year despite daily observations of sufficient intensity to identify all resident birds. Thus, surreptitious copulations around the time of laying could not possibly have accounted for the fertilization of all excluded eggs. (2) The five males that could not be excluded, and whose stored sperm is presumed to have accounted for fertilization of

excluded eggs, copulated repeatedly and cooperatively while paired to females that laid the excluded eggs. (3) Incubating males were excluded and (4) neighbouring males were excluded. (5) Males with whom the females attempted EPC were excluded, and (6) during 1,190 h of observation no successful EPCs (defined by cloacal contact) were observed between the females that laid these 11 eggs and any other males. Minimal estimates of the time that sperm was stored (the time from when a female and the nonexcluded mate last associated to the day before the excluded egg was laid) averaged 18.6 days for the five males described in (2) (range 3–31 days).

Previously, we found that early-arriving males had no greater reproductive success than late-arrivers¹⁴ thus we could not explain why older males preceded the arrival of yearlings to the breeding grounds and competed vigorously for early males¹¹. Here we show that males arriving and pairing early can gain an advantage, not through the greater success of clutches they incubate, but through additional fertilizations from the sperm stored by females from early-season matings. If it is assumed that older males are, on average, superior to younger males, this provides a 'passive' method for females to increase the probability of fertilization by higher quality males, rather than actively selecting superior males for EPCs³. In spotted sandpipers, EPFs resulting from stored sperm are an important source of variance in male reproductive success, giving early-pairing males higher reproductive success and greater confidence of paternity than later-pairing males. Late-arriving (younger) males face the

dilemma of not breeding, or breeding in a system in which their reproductive success might be reduced by EPFs resulting from stored sperm. This study further emphasizes that in all species with internal fertilization, and where females have several mates, an assessment of parentage using molecular techniques is vital to correctly determine reproductive success. □

Received 5 May; accepted 24 August 1992.

1. Jeffreys, A. J., Wilson, V. & Thein, S. L. *Nature* **316**, 76–79 (1985).
2. Gyllenstein, U. B., Jakobsson, S. & Temrin, H. *Nature* **343**, 168–170 (1990).
3. Kempenaers, B. *et al.* *Nature* **357**, 494–496 (1992).
4. Burke, T., Davies, N. B., Bruford, M. W. & Hatchwell, B. J. *Nature* **338**, 249–251 (1989).
5. Burke, T. & Bruford, M. W. *Nature* **327**, 149–152 (1987).
6. Birkhead, T. R., Burke, T., Zann, R., Hunter, F. M. & Krupa, A. P. *Behav. Ecol. Sociobiol.* **27**, 315–324 (1990).
7. Birkhead, T. R. *Adv. Study Behav.* **18**, 35–72 (1988).
8. Oring, L. W. & Lank, D. B. *Behav. Ecol. Sociobiol.* **10**, 185–191 (1982).
9. Oring, L. W., Lank, D. B. & Maxson, S. J. *Auk* **100**, 272–285 (1983).
10. Oring, L. W. & Lank, D. B. in *Ecological Aspects of Social Evolution* (eds Rubenstein, D. I. & Wrangham, R. W.) 21–42 (Princeton Univ. Press, Princeton, 1986).
11. Maxson, S. J. & Oring, L. W. *Behaviour* **74**, 200–263 (1980).
12. Oring, L. W. in *Avian Biology* Vol. 6 (eds Farner, D. S., King, J. R. & Parkes, K. C.) 1–92 (Academic, New York, 1982).
13. Oring, L. W. *Curr. Orn.* **3**, 309–351 (1986).
14. Oring, L. W., Reed, J. M., Colwell, M. A., Lank, D. B. & Maxson, S. J. *Behav. Ecol. Sociobiol.* **28**, 433–442 (1991).
15. Colwell, M. A. & Oring, L. W. *Anim. Behav.* **38**, 675–684 (1989).
16. Hatch, S. A. *Auk* **100**, 593–600 (1983).
17. Shugart, G. W. *Auk* **105**, 379–384 (1988).
18. Whitfield, D. P. *Behav. Ecol. Sociobiol.* **27**, 247–254 (1990).
19. Møller, A. P. & Birkhead, T. R. *Behaviour* **118**, 170–176 (1991).
20. Birkhead, T. R. & Møller, A. P. *Sperm Competition in Birds: Evolutionary Causes and Consequences* (Academic, London, 1992).
21. Møller, A. P. & Birkhead, T. R. *Biol. J. Linn. Soc. Lond.* **38**, 119–131 (1989).
22. Birkhead, T. R., Atkin, L. & Møller, A. P. *Behaviour* **101**, 101–138 (1987).
23. Burke, T. *Trends Ecol. Evol.* **4**, 139–144 (1989).
24. Rice, W. R. *Biometrics* **44**, 1–22 (1988).
25. Zenone, P. G., Sims, M. E. & Erickson, C. J. *Am. Nat.* **114**, 615–626 (1979).
26. Birkhead, T. R., Pellatt, J. & Hunter, F. M. *Nature* **334**, 60–62 (1988).
27. Westneat, D. F., Noon, W. A., Reeve, H. K. & Aquardo, C. F. *Nucleic Acids Res.* **16**, 4161 (1988).
28. Westneat, D. F. *Behav. Ecol. Sociobiol.* **27**, 67–76 (1990).
29. Lansman, R. A., Shade, R. O., Shapiro, J. F. & Avise, J. D. *J. molec. Evol.* **17**, 214–226 (1981).

ACKNOWLEDGEMENTS. We thank J. Alberico, J. Arp and S. Maxson for help in the field; K. Oring, G.-A. Maxson and Huddle's Resort for help in incubating eggs; P. Pickett, D. Hegge, E. Rave and K. Gibson for laboratory assistance; W. Eberhard, J. Rhymer and R. Wagner for reviewing the paper; and A. J. Jeffreys and Cellmark Diagnostics for permission to use the Jeffreys' 33.15 probe. This work was supported by the U.S. National Science Foundation.

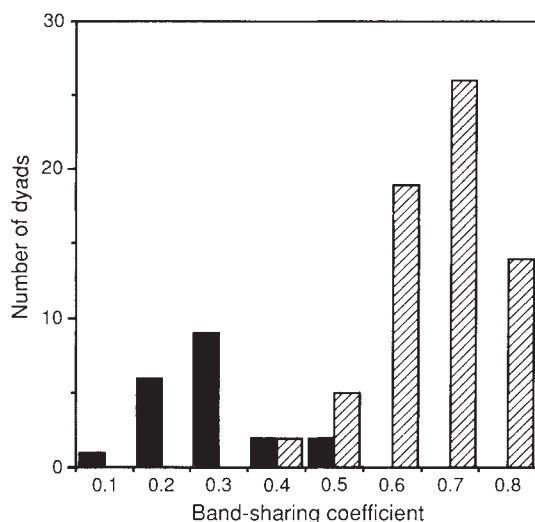


FIG. 2 Band-sharing coefficients (F in ref. 29) were calculated between all pairs of individuals within families (dark bars) and between offspring and suspected extra-pair males (striped bars). Shown here are distributions of F for all unexcluded father-offspring dyads (mean $F=0.66$, $s.d.=0.09$, $n=66$ dyads) and 20 random dyads (pairs of probably unrelated individuals; mean $F=0.29$, $s.d.=0.09$). Average band-sharing between the incubating male and the offspring excluded from it were similar to the value for unrelated males (mean $F=0.31$, $s.d.=0.08$, range, 0.16–0.43), whereas values of F for laying females were similar to those expected from first-order relatives (mean $F=0.55$, $s.d.=0.07$, range, 0.43–0.64). A single egg showed low band-sharing with the female ($F=0.31$), and high band-sharing with the male ($F=0.68$), suggesting a case of intraspecific egg parasitism. Thus, 11 of 12 exclusions were attributable to extra-pair fertilization; the remaining egg probably resulted from brood parasitism. Offspring profiles were also compared to profiles of males observed attempting copulations with the laying female, the female's previous mates, and males on adjacent territories. We excluded all but one of these males for 10 of the 11 exclusions attributable to EPF. Each non-excluded male was a previous mate of the female (Table 1). The single remaining offspring originated from a female's first clutch, where she had no previous mate. The value of F between these offspring and the non-excluded males averaged 0.62 ($s.d.=0.12$, $n=10$ dyads, range, 0.46–0.80). These data support assignment of excluded offspring to previous mates. Exclusions were confirmed by a second round of fingerprint analyses.

High outcrossing rates maintain male and hermaphrodite individuals in populations of the flowering plant *Datisca glomerata*

Peter Fritsch & Loren H. Rieseberg

Rancho Santa Ana Botanic Garden, 1500 North College Avenue, Claremont, California 91711-3101, USA

MODELS for the maintenance of androdioecy (the presence of male and hermaphrodite individuals in a breeding population) in plants predict that males must have a fertility at least double the male fertility of hermaphrodites in order to be maintained by selection^{1–3}. An even greater advantage is required in partially self-fertilizing populations^{1–3} as the gain in fitness through increased pollen production is least when few ovules are available for outcrossing. Because such stringent theoretical requirements make the evolutionary stability of this breeding system highly unlikely, functional androdioecy is thought to be rare in plants, and indeed the only documented instance occurs in populations of *Datisca glomerata* (Datisceae)⁴. As such, these populations provide a unique opportunity to test predictions concerning the evolution of androdioecy in plants. Here we report high outcrossing rates (65–92%) in two androdioecious populations of *D. glomerata* using random amplified polymorphic DNA markers. These

outcrossing rates, when analysed with respect to existing evidence concerning pollen production and inbreeding depression in this species, are sufficiently high to satisfy theoretical requirements for the maintenance of androdioecy.

A previous study⁵ of genetic variation in *D. glomerata* using isozyme electrophoresis revealed an inbreeding coefficient (F) of 0.617 ($P < 0.001$). This value of F suggests high rates of self-fertilization, whereas models for the maintenance of males in androdioecious populations require low rates. But as F is a product not only of self-fertilization, but of any type of non-random mating and selection, an estimate of cross-fertilization among genets (genetically distinct individuals, each arising from a single zygote) was needed to test theoretical models for the maintenance of androdioecy.

Outcrossing estimates can be generated by examining allelic variation over many loci in progeny arrays from plants allowed to pollinate naturally (open-pollination) and using a maximum-likelihood model to exclude progeny resulting from self-fertilization⁶⁻⁸. Initial investigations employing isozyme electrophoresis produced insufficient numbers of polymorphic loci with which to exclude progeny. Therefore, we have examined random amplified polymorphic DNA (RAPD)⁹ loci in open-pollinated progeny arrays of *D. glomerata* to obtain multilocus estimates of outcrossing.

The progeny of 12 hermaphrodite plants from an androdioecious population of *D. glomerata* 6 km south of Islip Saddle, Los Angeles County, California (Mt Islip⁴) were grown to a

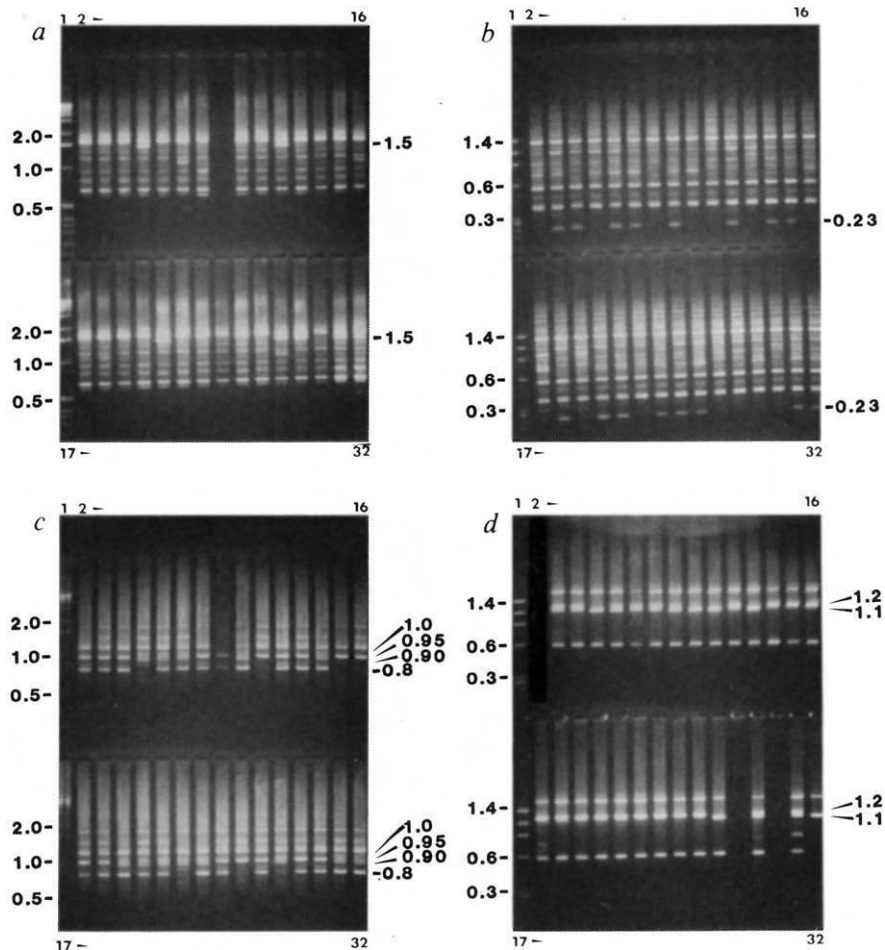
size sufficient for total DNA isolation¹⁰ from fresh leaves or whole seedlings. Maternal DNAs were isolated from dried green capsular tissue. Families were surveyed for genetic polymorphism by DNA amplification and electrophoresis¹¹ of selected individuals. From a survey of 340 arbitrary 10-base oligodeoxynucleotide primers (primers 201-500, University of British Columbia Biotechnology Laboratory (UBC); primer sets C and D, Operon Technologies (OP)), we inferred 12 polymorphic loci (Table 1) which we used to determine multilocus single-family and population outcrossing rates (Table 2 and Fig. 1a-c). The multilocus outcrossing rate, t_m , can range from 0 to 1 and when multiplied by 100 represents the estimated percentage of genet-to-genet cross-fertilizations within a population. We estimated t_m with a maximum likelihood algorithm¹² modified to allow inclusion of dominant markers. All dominant loci were considered diallelic, thus all individuals lacking a fragment were assumed to be homozygous recessive for the same allele.

For the Mt Islip population, t_m was 0.92 (s.e.m. = 0.03). The predicted frequency of male plants in androdioecious populations under conditions of equal male and female zygote fitness is given by

$$q = \frac{t - 2lv(t + i - ti)}{2iv(1-t)(1-l) + t(1+v-2lv)} \quad (1)$$

where t is outcrossing rate, l is relative (female/male) pollen fecundity, i is relative (inbred/outbred) fitness of seeds due to

FIG. 1 RAPD profiles of selected families used in the estimation of outcrossing rates in *D. glomerata*. *a*, Mt Islip families 9, 10 and 11 amplified with Operon primer C-19. Lanes 1 and 17, kb ladder (Gibco BRL); lanes 2-7, family 9 (partial array; maternal parent not shown); lanes 8-16, 18-20, family 10 (maternal parent, lane 8); lanes 21-32, family 11 (maternal parent, lane 21). Progeny in lanes 5, 12, 22, 28, 29, 31 and 32 can be excluded from self-fertilization on the basis of the presence of the 1.5 kb allele. *b*, Mt Islip family 2 amplified with UBC primer 287. Lanes 1 and 17, PhiX174 DNA-HaeIII digest marker (New England Biolabs); lane 2, maternal parent; lanes 3-16 and 18-32, progeny. Progeny in lanes 3, 4, 6, 7, 9, 12, 14, 15, 19, 21, 22, 24-26, 31 and 32 can be excluded from self-fertilization on the basis of the presence of the 0.23 kb allele. *c*, Mt Islip family 6 amplified with UBC primer 437. Lanes 1 and 17, 1 kb ladder (Gibco BRL); lane 2, maternal parent; lanes 3-16 and 18-32, progeny. The fragments at 0.90 kb and 0.95 kb represent codominantly inherited alleles at a single locus. The maternal parent is homozygous for the slow (0.95 kb) allele, thus excluding from self-fertilization the heterozygous progeny in lanes 5-7, 10, 14, 20-24, 28, 31 and 32; the fragment at 1.0 kb, which amplifies only when the progeny are heterozygous, is assumed to be artifactual. The dominant allele of another locus at 0.8 kb is heterozygous in the maternal parent, thus, genotypes of progeny with the 0.8 kb allele cannot be determined. *d*, Alder Creek families 1, 2 and 3 amplified with UBC primer 106. Lanes 1 and 17, PhiX174 DNA-HaeIII digest marker (New England Biolabs); lanes 2-13, family 1 (maternal parent not shown); lanes 14-16 and 18-26, family 2 (maternal parent, lane 14); lanes 27-32, family 3 (partial array; maternal parent, lane 27). The fragments at 1.1 kb and 1.2 kb represent codominantly inherited alleles at a single locus. The maternal parent of family 1 (not shown) is homozygous for the fast (1.1 kb) allele, thus excluding from self-fertilization the heterozygous progeny in lanes 3, 4, 6-9 and 12. The maternal parent of family 2 is homozygous for the slow (1.2 kb) allele, as are all



progeny, and therefore none can be excluded. The maternal parent of family 3 is homozygous for the fast allele, and therefore the heterozygous progeny in lanes 29 and 31 can be excluded.

TABLE 1 Polymorphic loci used to determine t_m in populations of *D. glomerata*

Locality/Primer	Locus	Allelic mobility (kb)
Mt Islip		
OP C-19	1	1.5
	2	1.2
UBC 287	3	0.75
	4	0.23
UBC 300	1	1.8
UBC 341	1	1.8
UBC 376	1	0.80
UBC 437*	1	0.90, 0.95
	3	0.8
UBC 474	1	1.1
	2	0.72
UBC 490	1	0.9
Alder Creek		
OP A-8*	1	0.92, 0.95
UBC 104*	1	1.36, 1.40
	2	1.0
	3	0.75
UBC 106*	1	1.1, 1.2
UBC 152	1	1.9
	2	1.7
	3	0.6
UBC 174	1	1.05
	2	0.7
UBC 287	1	1.6
	2	1.0
	3	0.75
UBC 341	1	1.8
UBC 409*	1	0.58, 0.59
UBC 413	1	0.8
UBC 437	2	0.85
UBC 474	2	0.72
	3	0.70

For each primer, amplification products representing polymorphic loci are ordered arbitrarily by size; the largest fragment is designated as locus 1. Codominance was confirmed in all observed cases by excising amplification products representing each suspected codominant allele from the gel exhibiting the RAPD profile, cleaving each with a series of restriction endonucleases, and observing congruent band profiles. OP, Operon Technologies; UBC, University of British Columbia; kb, kilobases.

* Codominant allelic expression.

inbreeding depression, and v is relative (female/male) viability of zygotes to reproductive maturity². Substituting our value of t_m for t , 0.26 for l (ref. 4), 0.86 for i (ref. 13; the estimate of i is based only on early phases of the life cycle, but in any case has little influence on q relative to t and l) and assuming $v = 1$ (no data), then $q = 0.28$. As $q > 0$, our estimate of t satisfies theoretical requirements for the maintenance of androdioecy in this population.

To confirm high outcrossing in androdioecious populations of *D. glomerata*, we estimated t_m for another androdioecious population 7 km south of Roundtop summit, Los Angeles County, California (Alder Creek⁴). From a survey of 163 primers (UBC 101–200, OP sets A and B, plus 23 robust primers from the Mt Islip survey), we inferred 19 polymorphic loci (Table 1) which were subsequently used to determine t_m (Table 2 and Fig. 1d). Only 3 of the 12 loci that were polymorphic in Mt Islip were detected as polymorphic in Alder Creek. In this population, t_m was 0.65 (s.e.m. = 0.05). Substituting t_m for t in equation (1) and keeping the other parameters the same, $q = 0.11$. Once again, outcrossing rates are sufficiently high to account for the maintenance of androdioecy.

The observed frequency of males in the two populations is $q = 0.24$ for Mt Islip and $q = 0.17$ for Alder Creek⁴. The Mt Islip value is 0.02 units lower than the q calculated using the lower

TABLE 2 Single family and population outcrossing estimates in *D. glomerata*

Locality/Family	N	$t_m \pm \text{s.e.m.}$
Mt Islip		
1	29	0.85 $\pm$ 0.08
2	29	0.86 $\pm$ 0.13
3	30	1.04 $\pm$ 0.20
4	28	1.10 $\pm$ 0.19
5	27	0.93 $\pm$ 0.31
6	31	1.01 $\pm$ 0.33
7	10	0.80 $\pm$ 0.33
8	9	1.00 $\pm$ 0.00
9	10	1.00 $\pm$ 0.00
10	11	0.78 $\pm$ 0.33
11	11	0.93 $\pm$ 0.54
12	11	0.56 $\pm$ 0.14
Population	236	0.92 $\pm$ 0.03
Alder Creek		
1	11	0.64 $\pm$ 0.25
2	11	0.28 $\pm$ 0.15
3	11	0.99 $\pm$ 0.55
4	11	1.00 $\pm$ 0.00
5	11	1.00 $\pm$ 0.00
6	11	0.84 $\pm$ 0.62
7	9	1.00 $\pm$ 0.00
Population	75	0.65 $\pm$ 0.05

N, number of progeny in each family; t_m , multilocus outcrossing rate estimate. Standard errors are each based on 100 bootstraps.

limit of the standard error for t_m ; the Alder Creek value is 0.02 units higher than the q calculated using the higher limit of the standard error for t_m . Nevertheless, observed and expected values of male frequency are surprisingly close, and indicate that at least some populations of *D. glomerata* have male frequencies in equilibrium with an androdioecious breeding system; apparently because hermaphrodite plants have only about one quarter of the male fertility of male plants.

We suggest that the differences in outcrossing rate observed between Mt Islip and Alder Creek are density dependent. Mt Islip, with higher values of q and t , is a compact population, with inflorescences from several individuals often adjacent or touching, whereas Alder Creek, with lower values of q and t , is a sparse, scattered population. This hypothesis will be tested by conducting detailed mapping studies in conjunction with paternity exclusion analysis for several androdioecious populations with density differences.

The extreme rarity of androdioecy in plants seems to be paralleled by the rarity of comparable breeding systems in animals. Mixed populations consisting of self-fertilizing hermaphrodites and males have been documented only in the shrimp genera *Eulimnadia*¹⁴ and *Triops*¹⁵, and probably occur in the barnacle genus *Iblaquadralvis* (P. Raimondi, personal communication). Comparison of the natural history of *D. glomerata* to these vastly different organisms may provide additional insights regarding the evolutionary circumstances under which androdioecy originates and is maintained.

In addition to confirming theoretical models for the maintenance of androdioecy, this study demonstrates the utility of RAPD markers in outcrossing rate estimations, especially when isozymes fail to provide sufficient polymorphism. For future outcrossing studies using RAPD technology, we suggest the development of codominant loci using restriction endonuclease cleavage of products amplified with RAPD primers, because dominant markers, which constitute about 95% of all RAPD markers obtained^{9,11}, yield little information if present in the maternal parent. This will result in the need for fewer numbers of polymorphic loci for reliable multilocus outcrossing

estimates. Codominant markers are currently being developed for populations of *D. glomerata*. □

Received 15 July; accepted 1 September 1992.

1. Charlesworth, D. *Biol. J. Linn. Soc. Lond.* **23**, 333–348 (1984).
2. Lloyd, D. G. *Genetica* **45**, 325–339 (1975).
3. Charlesworth, B. & Charlesworth, D. *Am. Nat.* **112**, 975–997 (1978).
4. Liston, A., Rieseberg, L. H. & Elias, T. S. *Nature* **343**, 641–642 (1989).
5. Liston, A., Rieseberg, L. H. & Elias, T. S. *Aliso* **12**, 525–542 (1989).
6. Ritland, K. & Jain, S. *Heredity* **47**, 35–52 (1981).
7. Brown, A. H. D. & Allard, R. W. *Genetics* **66**, 133–145 (1970).
8. Clegg, M. T., Kahler, A. L. & Allard, R. W. *Genetics* **89**, 765–792 (1978).
9. Williams, J. K. G., Kubelik, A. R., Livak, K. J., Rafalski, J. A. & Tingey, S. V. *Nucleic Acids Res.* **18**, 6531–6535 (1990).
10. Rieseberg, L. H., Hanson, M. A. & Philbrick, C. T. *Syst. Bot.* **17**, 324–336 (1992).
11. Rieseberg, L. H., Choi, H., Chan, R. & Spore, C. D. *Heredity* (in the press).
12. Ritland, K. *J. Hered.* **81**, 235–237 (1990).
13. Rieseberg, L. H., Philbrick, C. T., Paek, P. E., Hanson, M. A. & Fritsch, P. *Am. J. Bot.* (in the press).
14. Sassaman, C. *Bull. mar. Sci.* **45**, 425–432 (1989).
15. Sassaman, C. *Hydrobiologia* **212**, 169–179 (1991).

ACKNOWLEDGEMENTS. We thank R. Chan, H. Choy, M. A. Hanson, C. T. Philbrick, and C. D. Spore. This work was supported by a grant from the NSF (L.H.R.).

Dissociation between mental imagery and object recognition in a brain-damaged patient

Marlene Behrmann*†‡, Gordon Winocur*†§
& Morris Moscovitch*†

* Rotman Research Institute of Baycrest Centre, 3560 Bathurst Street, Toronto, Ontario M6A 2E1, Canada

† Department of Psychology, Erindale College, University of Toronto, Mississauga, Ontario L5L 1C6, Canada

‡ Faculty of Medicine, University of Toronto, 88 College Street, Toronto, Ontario M5G 1L4, Canada

§ Department of Psychology, Trent University, Peterborough, Ontario K9J 7B8, Canada

VISUAL imagery is the creation of mental representations that share many features with veridical visual percepts. Studies of normal and brain-damaged people reinforce the view that visual imagery and visual perception are mediated by a common neural substrate and activate the same representations^{1–4}. Thus, brain-damaged patients with intact vision who have an impairment in perception should have impaired visual imagery. Here we present evidence to the contrary from a patient with severely impaired object recognition (visual object agnosia) but with normal mental imagery. He draws objects in considerable detail from memory and uses information derived from mental images in a variety of tasks. In contrast, he cannot identify visually presented objects, even those he has drawn himself. He has normal visual acuity and intact perception of equally complex material in other domains. We conclude that rich internal representations can be activated to support visual imagery even when they cannot support visually mediated perception of objects.

C.K. is a 33-year-old, right-handed man who emigrated from England to Canada in 1980. C.K. presented with a severe deficit in object perception following a closed head injury sustained in a motor vehicle accident in 1988. Premorbidly, C.K. was enrolled in a Master's degree in history which he recently completed with special educational aids. C.K. has a partial left homonymous hemianopia and his visual acuity is 6/7.5 with corrective lenses. No focal lesion is seen on either computed tomography scan (December 1991) or magnetic resonance imaging scan (June 1992) although there is a suggestion of thinning of the occipital lobes bilaterally. The right lateral ventricle was slightly larger than the left but this difference was considered to be within normal limits.

C.K. was able to identify only 18/60 black and white line drawings from a standardized test on which age- and schooling-

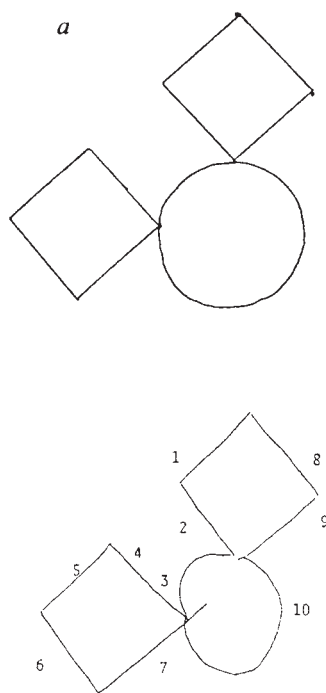
matched normal performance is 57/60 (ref. 5). For example, in the Boston Naming test: an asparagus was identified as "a rose twig with thorns", a tennis racquet as "a fencer's mask" and a dart as "a feather duster". His errors indicate that, like some other visual agnostic patients^{6,7}, he proceeds in a piecemeal fashion, reconstructing elements of the stimulus and then inferring the object's identity rather than recognizing the object as a meaningful whole. When presented with three-dimensional common objects, he correctly named 11/23 (normal performance 23/23), calling a saw "a knife", pliers "a clothes peg" and a padlock "an earring". With visually presented letters, he was unable to identify any; however, he could identify all of them with perfect accuracy when permitted to trace them. It is unlikely that the deficit depends on the size of the stimulus. C.K. was impaired when stimuli, both letters and pictures, ranged in size from a few millimetres to several centimetres.

Because C.K. named all the objects when he could palpate them with his eyes closed, the object perception deficit cannot be explained by a difficulty in recalling the names of the objects nor can it be attributed to a loss of knowledge of what the objects are. He can provide detailed verbal definitions for objects: for example, he defined a duck as "an animal, marine life, with webbed feet and a bill". The deficit in object recognition also cannot be explained by a global perceptual impairment. Remarkably, C.K.'s ability to identify faces is preserved. He easily recognizes famous people from photographs, even those that make identification difficult. Also, he scored 49 out of 54 (mean for normal males is 46) on a difficult standardized face recognition test that includes matching faces photographed in front and three-quarter view, or under poor lighting conditions⁸. His preserved face recognition suggests that he can adequately perceive and identify stimuli that are at least as complex as the objects he is unable to recognize.

Despite his profound deficit, C.K. can copy figures accurately although he does so slavishly and segmentally. For example, he copies geometric configurations correctly (Fig. 1a), but draws the segments in an unusual order, suggesting that he cannot appreciate the identity of the objects. Similarly, he can copy text but does so slowly and concretely (Fig. 1b). Together, these symptoms are classical features of visual object agnosia.

In contrast to C.K.'s perceptual deficit, his ability to form mental images is normal. C.K. is particularly good at drawing and he took art classes in high school. He can draw from memory complex objects with rich detail (Fig. 1c). He was even able to provide an accurate rendition of all but one (a seahorse) of the 30 items that he failed to recognize on the Boston Naming test⁵. When shown his own drawings on a subsequent occasion, he was unable to identify any of them. This suggests that despite his severe perceptual deficit, he has retained a long-term representation of the objects which he can then use for drawing⁹.

To assess further his visual imagery, we administered six standard tests that have been used with normal and brain-damaged subjects as tests of mental imagery. On a letter task^{10,11}, C.K. was asked to imagine an upper-case letter and to judge whether it has any curved lines (C versus L) or to imagine a lower-case letter and to decide whether the letter has a line ascending above or descending below the body (b versus p). He performed perfectly, scoring 26/26 on each test as do normal subjects. He also scored perfectly (20/20) on a colour imagery task¹², naming the characteristic colour of common items whose colours are not verbally associated with the item (for example, a football, the inside of a cantaloupe). On a size comparison task^{12,13}, in which he judged which of two similar-sized objects was larger, C.K. was able to make the correct judgements in all 16 pairs (for example, a popsicle and a pack of cigarettes, a vase and a hydrant). He scored 20/20 on an animal tails task^{12,14} in which he judged whether the animal has a long or short tail relative to its body size (for example, kangaroo, pig) and 20/20 on an animal ears test¹⁴ in which he decided whether the animal has floppy or upright ears (for example, doberman, dachshund).



b

Pack my box with five dozen jugs of liquid veneer.

Pack my box with five dozen jugs of liquid veneer.

Pack my box with five dozen jugs of liquid veneer.



FIG. 1 Examples of C.K.'s performance on various tasks. *a*, Example of C.K.'s copying of geometric configurations. The numbers indicate the sequential order of the strokes. *b*, C.K.'s copying (middle) and spontaneous writing (bottom) of a target sentence (top). Note the difference between the copied letter 'a' which is identical to that in the target stimulus and the letter 'a' produced in spontaneous writing to dictation. *c*, Examples of C.K.'s spontaneous drawings of items from memory: a map of England with 'X' marking his birthplace and an electric guitar.

C.K. also performed flawlessly on mental imagery tasks on which his knowledge of distance and spatial layout (rather than an object's appearance) was tested. For example, on a spatial estimation test in which he was given the names of three places (Manchester, Leeds, Birmingham) and had to judge which two are closest together, he scored 20/20. On a clock test¹⁵, he could decide with perfect accuracy (30/30) whether the angle subtended by the two hands of a clock is larger or smaller than 90 degrees (such as 6.20, 1.35)¹⁹. Finally, C.K. was given the Brooks letter test¹⁶, in which he was instructed to imagine a large block-capital letter (E or F) laid out on the ground and to imagine walking clockwise along the outside of the letter, starting on the bottom left corner. He was to report his route and to indicate whether he made right or left turns every time he reached a corner. C.K. indicated only one incorrect turn on this task which was well within normal limits of performance.

C.K. thus has excellent mental imagery despite poor recognition of objects. His perceptual deficit is of central origin as indicated by his intact visual acuity and normal perception of complex stimuli (faces). That he can define objects verbally and identify them by touch suggests that he has retained knowledge of objects he cannot perceive. No other case of such a clear dissociation between impaired object perception and intact imagery has been reported so far.

C.K. complements the reported cases who have intact perception together with impaired mental imagery^{17,18}, thus completing both aspects of the double dissociation. Together, these data challenge the prevailing view that a deficit in perception of central origin arises from a loss of the mental representations and the neural structures that mediate them. Our results suggest two alternative hypotheses. One hypothesis is that there are separate underlying representations in long-term memory, one set used for generating mental images and a second set used for perceiving objects. An alternative hypothesis is that there is

a single set of higher-level visual representations with separate routes of access for imagery and for perception. By this view, the dissociation between mental imagery and object perception can arise because access to shared representations cannot be gained either by an internal route for imagery or by an external one for perception. A variant of this view is that there are impairments in intermediate level processes involved in constructing a common representation. Either variant of the latter view is preferred because it provides a unifying and parsimonious account of both manifestations of the double dissociation. *Note added in proof:* Since submission of this letter, a visual agnostic (M.D.) has been described²⁰ who also shows some preserved imagery despite his perception deficit. □

Received 6 May; accepted 7 August 1992.

1. Farah, M. J. *Psychol. Rev.* **95**, 307-317 (1988).
2. Finke, R. A. *Psychol. Bull.* **98**, 236-259 (1985).
3. Farah, M. J. et al. *J. exp. Psychol.* **117**, 248-257 (1988).
4. Farah, M. J. *Cognition* **18**, 245-251 (1984).
5. Kaplan, E. et al. *Boston Naming Test* (Lea and Febiger, Philadelphia, 1983).
6. Riddoch, M. J. & Humphreys, G. W. *Brain* **110**, 1431-1462 (1987).
7. Gralet, J. M. et al. *Cognitive Neuropsychol.* **7**, 275-310 (1990).
8. Benson, A. L., deS. Hamsher, N., Varney, N. & Spreen, O. *Facial Recognition* (Oxford Univ. Press, New York, 1983).
9. van Sommers, P. *Cognitive Neuropsychol.* **6**, 117-164 (1989).
10. Kosslyn, S. M. et al. *J. exp. Psychol.* **114**, 311-341 (1985).
11. Sergent, J. *J. exp. Psychol. Hum. Percept. Perform.* **15**, 170-178 (1989).
12. Farah, M. J. et al. *Cognitive Psychol.* **20**, 439-462 (1988).
13. Kosslyn, S. M. et al. *J. exp. Psychol.* **106**, 341-376 (1977).
14. Kosslyn, S. M. *Cognitive Psychol.* **7**, 341-370 (1975).
15. Corballis, M. & Sergent, J. *Neuropsychologia* **26**, 12-26 (1988).
16. Brooks, L. R. *Can. J. Psychol.* **22**, 349-368 (1968).
17. Riddoch, M. J. *Cognitive Neuropsychol.* **7**, 249-273 (1990).
18. Farah, M. J. et al. *Brain Cogn.* **8**, 147-164 (1988).
19. Craik, F. I. M. & Dirks, E. *Psychol. Aging* **7**, (in the press).
20. Jankowiak, J., Kinsbourne, M., Shalev, R. & Bachman, D. L. *J. cogn. Neurosci.* **4**, 119-131 (1992).

ACKNOWLEDGEMENTS. This research was supported by grants from the Medical Research Council of Canada and by an Ontario Mental Health Foundation Research Associateship and a Medical Research Council Scholarship to M.M. and to M.B., respectively.

Hippocampus-dependent learning facilitated by a monoclonal antibody or D-cycloserine

Lucien T. Thompson*, Joseph R. Moskal*†
& John F. Disterhoft*

* Department of Cell, Molecular and Structural Biology,
Northwestern University Medical School, 303 East Chicago Avenue,
Chicago, Illinois 60611-3008, USA

† The Chicago Institute for Neurosurgery and Neuroresearch,
428 West Deming Place, Chicago, Illinois 60614, USA

PERSISTENT neuronal plasticity, including that observed at some hippocampal synapses, requires N-methyl-D-aspartate (NMDA)-mediated transmission. NMDA receptor activation may be necessary for hippocampus-dependent learning as antagonists block acquisition in many such tasks. The behavioural effects of NMDA agonists are less well defined. We have shown that a monoclonal antibody (B6B21) displaced [³H]-glycine that was bound specifically to the NMDA receptor, and enhanced the opening of its integral cation channel in a glycine-like fashion, effects that were competitively antagonized by 7-chlorokynurenic acid¹. B6B21 also enhanced long-term potentiation in hippocampal slices¹. We report here that intraventricular infusions of B6B21 significantly enhances acquisition rates in hippocampus-dependent trace eye blink conditioning in rabbits, halving the number of trials required to reach a criterion of 80% conditioned responses. Peripheral injections of D-cycloserine, a partial agonist of the glycine site on the NMDA receptor which crosses the blood-brain barrier, also doubles rabbits' learning rates. Pseudoconditioning control experiments indicated a lack of nonspecific behavioural sensitization effects. Our data suggest that enhanced activation of the glycine coagonist site on the NMDA receptor/channel complex facilitates one form of associative learning and may be used in other learning tasks.

Our studies were designed to determine whether compounds with glycine-like activity *in vitro* can affect associative learning *in vivo*. Partial agonists were used to avoid induction of seizures and other pathological manifestations of excessive activation of NMDA receptors². A hippocampus-dependent learning task was used (that is, one abolished by hippocampal ablation³). The hippocampus is particularly rich in NMDA receptors⁴. Long-term potentiation of Schaffer collateral or perforant-path synap-

ses is blocked by competitive antagonists of the NMDA binding site⁵ or of its strychnine-insensitive glycine coagonist site⁶. Intracellular recordings from neocortical pyramidal cells indicate that exogenous glycine can enhance single NMDA-mediated excitatory postsynaptic potentials (e.p.s.ps)⁷. Long-term potentiation and other forms of neuronal plasticity may share common synaptic or subcellular mechanisms with associative learning⁸. NMDA antagonists (both competitive and non-competitive) inhibit acquisition in behavioural learning tasks at doses that do not impair performance^{9,10}, but few have demonstrated any beneficial effects of NMDA or glycine-site agonists. Most previous agonist studies have used simple tasks that are affected non-selectively by a wide range of neurotransmitter systems¹¹.

In the first set of experiments, lateral ventricular guide canulae were surgically implanted bilaterally in rabbits, which were then fitted with restraining head bolts. After recovery and environmental habituation, rabbits were trained in pairs in separate, darkened and sound-attenuated chambers. Training continued until the criterion level of 80% of paired stimulus presentations resulted in the conditioned response (80% CR).

Intraventricular infusion of B6B21 greatly accelerated acquisition of trace conditioning as compared with either IgG or artificial cerebrospinal fluid control treatments ($F_{2,5} = 15.14$, $P < 0.0003$; see Fig. 1a). In every case, B6B21-treated animals learned much more rapidly than their paired controls. Learning-curve slopes did not differ significantly between the control groups (Scheffe F -test, $F = 0.068$; $P > 0.5$). B6B21 treatment resulted in a 45% reduction in the number of trials required to reach 80% conditioned responses (one-way analysis of variance, $F_{2,15} = 5.601$; $P < 0.01$) and reductions in other operational measures of learning (Fig. 1b). No effects of B6B21 treatment on the unconditioned response amplitude during conditioning were observed ($P < 0.68$).

After conditioning, extinction with conditioned stimulus presentation alone was carried out for an additional three days. All rabbits returned to a low level of conditioned response performance at about the same rate, and there was no effect of B6B21 treatment on the percentage of conditioned responses on the three days of testing ($F_{1,12} = 0.51$; $P > 0.45$). But a more subtle result of B6B21 treatment was observed when extinction was analysed in 10-trial blocks. Rabbits commonly exhibit a phenomenon (shown by controls) termed 'spontaneous recovery'¹², in which the percentage of conditioned responses in early blocks of trials on each successive day return briefly to levels near 'criterion' (80% CR) performance, followed by a

TABLE 1 Results of pseudoconditioning experiments

Treatment group	Unpaired CS trials				Unpaired US trials			
	Responses to CS (CR)		Responses after trace interval no US presented (false UR)		Responses with no CS presented before US onset (false CR)		Responses to US (UR)	
	Responses (Per cent of trials)	Response amplitude (mV)	Responses (Per cent of trials)	Response amplitude (mV)	Responses (Per cent of trials)	Response amplitude (mV)	Responses (Per cent of trials)	Response amplitude (mV)
B6B21	11.0 ± 4.5	2,705 ± 810	8.3 ± 4.1	2,247 ± 428	6.3 ± 3.2	2,408 ± 497	99.2 ± 6.4	3,290 ± 798
Control	15.2 ± 3.6	2,716 ± 613	9.8 ± 5.2	2,291 ± 511	5.9 ± 2.8	2,372 ± 658	96.1 ± 8.3	3,244 ± 877
D-Cycloserine	11.4 ± 6.8	2,522 ± 708	9.1 ± 3.6	2,193 ± 693	7.9 ± 4.2	2,356 ± 528	97.2 ± 3.8	2,995 ± 254
Control	16.0 ± 10.2	2,344 ± 344	7.3 ± 3.7	2,310 ± 648	8.7 ± 2.9	2,513 ± 502	97.5 ± 4.2	2,878 ± 260

CS, US, conditioned and unconditioned stimulus, respectively; CR, UR, conditioned and unconditioned response, respectively. Experiments were done with different groups of rabbits (4 per group). The nictitating membrane response was measured with an optical detector as in earlier studies (refs 3, 24, 36). Treatment with glycine partial agonists had no non-associative effects on eye-blink responses (data shown are mean ± s.e.m.). The same stimuli used in conditioning were presented unpaired at pseudorandom intervals 80 times each per session for 10 daily sessions. Alterations in unconditioned nictitating membrane responses (far right columns) would be indicative of nonspecific motor effects that might alter performance. All unconditioned response measurements including number of responses, response amplitude, and response latency (not shown) were unaffected by treatment with B6B21 or D-cycloserine (2 factor repeated measures ANOVA, $P > 0.43$). Changes in behavioural responses to the unpaired tone (far left columns) would be indicative of altered sensitivity to the tone, and might alter the salience of the CS when paired with a US. No measurable changes in responses to the unpaired tone were observed ($P > 0.38$). Nonspecific changes in behavioural excitability would be indicated by alterations in spontaneous eye-blink responses during periods when no stimuli were presented (including those shown in the middle four columns). Again, no significant differences in these measures were observed ($P > 0.59$).

rapid decline in response to the conditioned stimulus alone. B6B21-treated rabbits failed to recover spontaneously (Fig. 2a). This may indicate a more robust 'memory' of the previous day's experience, though the overall impact on extinction, a process that occurs very rapidly, was minimal.

Other groups of rabbits underwent pseudoconditioning with unpaired stimuli presented for equal numbers of trials for ten days, to test whether treatment with B6B21 produced nonspecific sensitization of responses to either the unpaired tone conditioned stimulus or the airpuff unconditioned stimulus. No differences were found in comparisons between B6B21 and control treated rabbits during pseudoconditioning (Table 1). Antibody B6B21 facilitated learning without altering response rates to the unpaired conditioned or the unconditioned stimulus, and without changing the behavioural characteristics of the small number of eye-blink responses to the unpaired tone or of the consistent and large number of responses to the air-puff.

A second series of studies was carried out using D-cycloserine (D-4-amino-3-isoxazolidone) a rigid analogue of D-alanine which acts as a partial agonist of the glycine coagonist site on the NMDA receptor¹³. It readily crosses the blood-brain barrier and is excreted unmetabolized¹⁴. As a partial agonist competing with glycine to bind at the glycine site on the NMDA receptor, D-cycloserine is less effective than the native agonist, yet still demonstrates reasonable specificity and high affinity for the receptor¹⁵.

For these studies, rabbits were surgically fitted with restraining headbolts. After recovery and environmental adaptation, pairs were trained using the same trace eye-blink conditioning task. Intramuscular injections of D-cycloserine (6 mg per kg body mass, given daily immediately before training) greatly acceler-

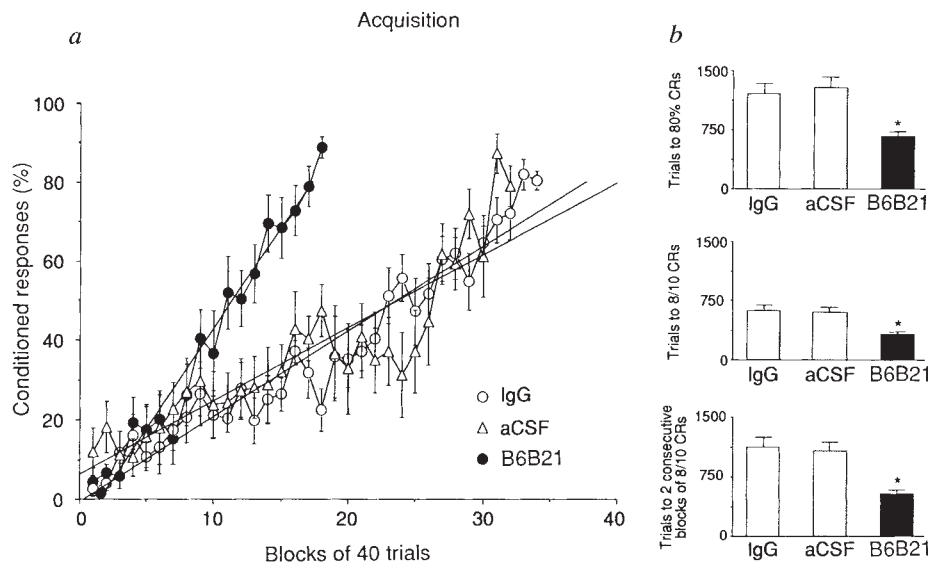
ated response acquisition. These rabbits exhibited learning curve slopes nearly twice as steep as those for controls (Fig. 3a; $F_{1,12} = 6.59$, $P < 0.02$) and the number of trials required to attain a level of performance of 80% CRs was reduced by 43% as compared with controls (one-way analysis of variance, $F_{1,12} = 6.04$, $P < 0.03$; Fig. 3b). Treatment with D-cycloserine had no effect on unconditioned response amplitude in conditioning ($P > 0.79$). Preliminary trials with a lower dose of D-cycloserine (3.0 mg per kg body mass) indicated a similar, though less effective enhancement of learning.

Treatment with D-cycloserine did not alter the percentage of CRs observed over three days of extinction of learned responses ($F_{2,15} = 0.002$; $P > 0.99$). But as in the earlier study, D-cycloserine-treated rabbits failed to recover spontaneously between successive days of extinction (Fig. 2b), demonstrating more uniform progressive extinction than controls. In pseudoconditioning studies, D-cycloserine-treated rabbits showed no evidence of sensitization or motor effects compared with saline-treated controls ($P > 0.38$; Table 1), which is similar to the results obtained with B6B21.

B6B21 and D-cycloserine greatly accelerated acquisition of trace eye-blink conditioning. Both act as partial agonists of the glycine site^{1,13,15}, and both facilitate NMDA-receptor mediated postsynaptic events^{1,16}. Although these molecules have quite different chemical structures, the epitope recognized by a monoclonal antibody may be equivalent to 3–6 amino acids¹⁷, thus it is probable that the glycine-like actions of both agents are mediated by interactions at the same recognition site. B6B21 competes for binding, in extensively washed hippocampal membrane preparations, with both the native agonist and with 7-chlorokynurenic acid, a competitive antagonist¹. B6B21

FIG. 1 Trace eye-blink conditioning was facilitated by daily intraventricular infusions of the monoclonal antibody B6B21. **a**, Mean learning curves ($\pm$ s.e.m.) and least-squares fitted lines for each curve are shown, expressed as the percentage of conditioned responses per block of 40 training trials. Using the linear interpolation algorithms of Igor (WaveMetrics, Lake Oswego, Oregon), each curve was normalized to the mean number of trials required to reach the criterion (80% CR) level for that group, so that qualitative summaries of learning rates for animals requiring different numbers of trials to reach criterion could be made. For statistical purposes, the slopes of individual non-normalized learning curves were compared, indicating that B6B21 treated rabbits learned significantly faster ($P < 0.003$) than rabbits in either control group (whose learning curves overlapped). **b**, Comparisons of the number of trials required to reach the criterion level, number of trials to reach a simpler criterion of 8/10 conditioned responses, and number of trials to gain two consecutive blocks of 8/10 conditioned responses all indicated significant facilitation of learning in the trace conditioning task by B6B21 ($P < 0.01$)*.

METHODS. Monoclonal antibodies were made against freshly dissected 5 day postnatal rat dentate gyri and purified using Protein A-Sepharose chromatography^{1,35}. Eluates were immediately neutralized, dialysed, and concentrated. Concentrated antibody (typically 1–2 ml) was again dialysed against two changes (each 2 l) of HEPES buffer to remove endogenous glycine contamination. The antibodies were sterilized by filtration through a 0.22- μ m filter and frozen at -80°C in small aliquots. All antibody injections were given under sterile conditions. Neither SDS-polyacrylamide gel electrophoresis nor low molecular mass high-performance liquid chromatography (HPLC) of purified antibodies revealed any glycine-like contaminants (our unpublished data). Specifically, no glycine or other α -amino acid contaminants were detected either in fresh samples of the antibody or in aliquots of the solutions infused intraventricularly, even when samples were



subjected to several freeze-thaw cycles (minimal HPLC detection threshold of 0.1 pmol per injection). Commercial IgG antibodies were used which had no glycine-like activity, and gave similar negative results from HPLC. Cannulated rabbits were always paired with an animal from one of the other two groups, with six animals used in each group. For 5 min immediately before each day's training, rabbits simultaneously received 5 μ l infusions in each ventricle of either B6B21 (1 μ g μ l⁻¹) suspended in artificial cerebral spinal fluid (aCSF; composition, in mM: 124 NaCl; 26 NaHCO₃; 3 KCl; 2.4 CaCl₂; 1.3 MgSO₄; 1.24 NaH₂O₄; 10 D-glucose (pH 7.4)), of aCSF alone, or of mouse IgG (1 μ g μ l⁻¹ Sigma) in aCSF at a rate of 1 μ l ventricle⁻¹ min⁻¹. The treatment received by the subjects was unknown to the trainer. Trace nictitating membrane conditioning (using a 100 ms duration, 6 kHz, 85 db binaural tone conditioned stimulus (CS) followed, after a 500 ms inter-stimulus trace interval by a 150 msec corneal air-puff unconditioned stimulus (US)) began immediately after infusion with 80 trials per day. For details of the protocols used, see refs 3, 36).

enhances NMDA-stimulated noradrenaline release in hippocampal slices, an effect also competitively antagonized by 7-chlorokynurenic acid (P. E. Potter and J. R. Moskal, unpublished observations). *In vitro*, both B6B21 (ref. 1) and D-cycloserine^{13,15,17} functionally mimic the effects of glycine and other α -amino acids, with demonstrable specificity and partial agonist efficacy. Unfortunately, many antagonists (including HA-966 (ref. 18) and various kynurenines, such as 5,7-dichlorokynurenic acid; T. Lanthorn, personal communication) are extremely toxic *in vivo*, confounding efforts to demonstrate adverse effects of glycine site antagonists on learning or other functions. Although such demonstrations are lacking in *in vivo* preparations, evidence from *in vitro* studies are consistent with the working hypotheses presented here^{1,6,13,15,16}.

Our data support the hypothesis that glycine-like partial agonists enhance some learning processes^{11,19,20} and suggest that, under normal conditions, NMDA receptor glycine sites are not saturated. The mechanisms involved in the uptake, release and buffering of glycine in extracellular spaces within the central nervous system are unclear, but our data further suggest that activation of NMDA receptors *in vivo* can be enhanced exogenously. Treatment with partial agonists like B6B21 or D-cycloserine could moderately enhance submaximal postsynaptic events or recruit additional synapses. Because the net change would yield greater NMDA-mediated responses throughout the learning process, the rate of learning would be accelerated. This interpretation is consistent with the facilitation of hippocampal long-term potentiation by low concentrations of glycine-site agonists²¹ or of partial agonists including B6B21 (ref. 1). Persistent postsynaptic changes in the excitability of hippocampal pyramidal neurons are seen after delay²² and trace²³ eye-blink conditioning. Although direct evidence linking behavioural changes to specific synaptic mechanisms is scarce, numerous authors have postulated the involvement of NMDA receptor-mediated plasticity in the induction, if not the expression, of learning^{8,9,10,16,19,20}.

The hippocampus may be an important locus of the behavioural effects of the glycine-like partial agonists reported here. Our observations that both B6B21 and D-cycloserine produce similar facilitation in this hippocampus-dependent task is consistent with this hypothesis, but does not eliminate the possibility that other sites of action may also be involved. Trace conditioning requires the formation of a short-term 'memory trace' of the conditioned stimulus to bridge the interstimulus interval between conditioned and unconditioned stimulus to successfully form an association. The hippocampus may serve this and other mnemonic functions^{3,24,26}. Acquisition of the 500-ms trace eye-blink conditioning task used here is blocked by hippocampal lesions in rabbits^{3,24}. When B6B21 was injected into the lateral ventricles, this large glycoprotein (M_r 150K) was unlikely to penetrate distal structures within a reasonable time for action^{27,28}. But the ventricular system gave relatively direct access to the hippocampus. Similarly, although D-cycloserine given peripherally reaches many brain areas, NMDA receptors are most abundant in the hippocampus⁴. Though hippocampal function is not an absolute requirement for all forms of learning, enhancement of hippocampal function may be of benefit even in tasks that are not dependent on the hippocampus. The effects of glycine-site partial agonists on the persistence of memory are currently unknown, as are the critical periods when treatment is effective both on learning and memory. These and other questions warrant further study.

The monoclonal antibody B6B21 may be valuable in further characterizing the structure and function of the NMDA receptor-ionophore complex. Moreover, by obtaining sequence information on the variable region of B6B21, it may be possible to generate biologically active peptides or mimetics of greater potency and specificity, and thus greater clinical relevance^{27,29}. As eye-blink conditioning tasks in rabbits and humans have important behavioural and neurobiological parallels^{30,31}, our experiments should be relevant to the role of the NMDA receptor glycine coagonist site in some forms of human learning.

FIG. 2 Glycine site partial agonists had more subtle effects upon extinction than upon acquisition of learned responses. *a*, Conditioned eye-blinks by control and B6B21-treated rabbits seemed to extinguish at nearly identical rates, with subjects in all groups returning to naive performance levels within 3 daily sessions (80 trials per session) of CS presentation alone. An important difference is that B6B21-treated rabbits failed to show the spontaneous recovery¹² of CRs shown by control animals (arrows). After one day of extinction, in the next block of 10 trials, controls exhibited a recovery followed by a rapid decrease in the number of CRs, to a level that was successively lower than that observed on the previous day. Extinction for controls was not a simple linear process, but instead entailed behavioural curves resembling 'saw-tooth' functions. Thus, for the controls on the second and third day of extinction the number of CRs in the first block of 10 trials was significantly higher than for the last block of 10 trials in that same session (extinction: second day, one-way paired *t*-test, $t=2.84$, $P<0.02$; third day, $t=3.84$, $P<0.01$) and also compared to the last block in the preceding session (spontaneous recovery: second day, $t=3.13$, $P<0.02$; third day, $t=2.38$, $P<0.03$). But for B6B21-treated rabbits only the former comparison was true (second day, $t=3.31$, $P<0.01$; third day, $t=2.91$, $P<0.02$; extinction alone occurred). B6B21-treated rabbits exhibited no spontaneous recovery ($P>0.14$), but instead exhibited extinction curves that were linear functions of the number of non-paired trials. *b*, D-cycloserine treatment also resulted in a loss of spontaneous recovery during extinction of conditioned eye-blink responses (arrows). Control rabbits exhibited both extinction (second day, $t=2.47$, $P<0.03$; third day, $t=2.89$, $P<0.01$) and spontaneous recovery (second day, $t=2.91$, $P<0.01$; third day, $t=2.52$, $P<0.02$) for CS-alone trials whereas D-cycloserine-treated rabbits exhibited extinction alone (second day, $t=3.29$, $P<0.01$; third day, $t=3.36$, $P<0.01$) with no spontaneous recovery ($P>0.18$). Again, first order functions described the extinction curves of D-cycloserine-treated but not those of control rabbits.

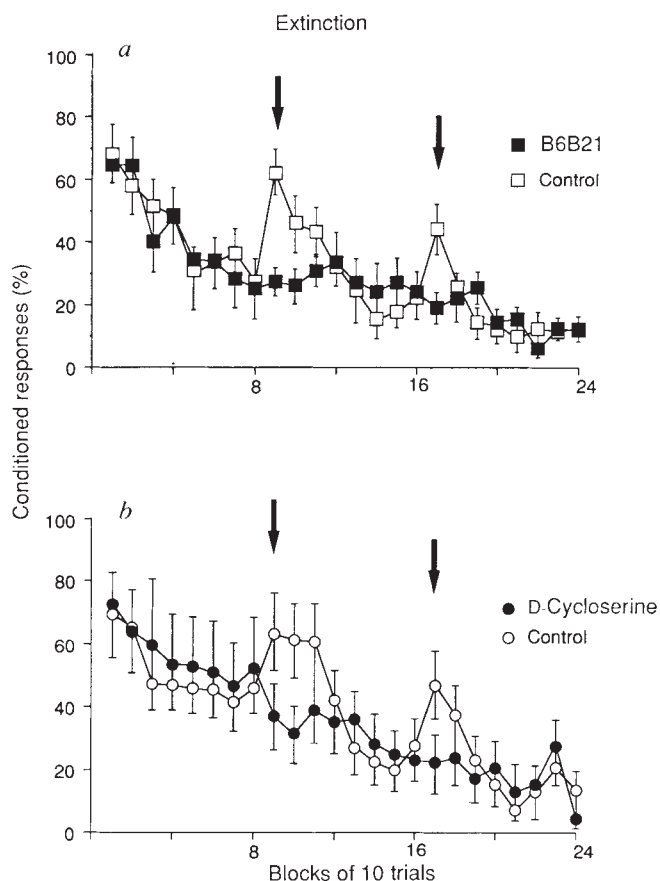
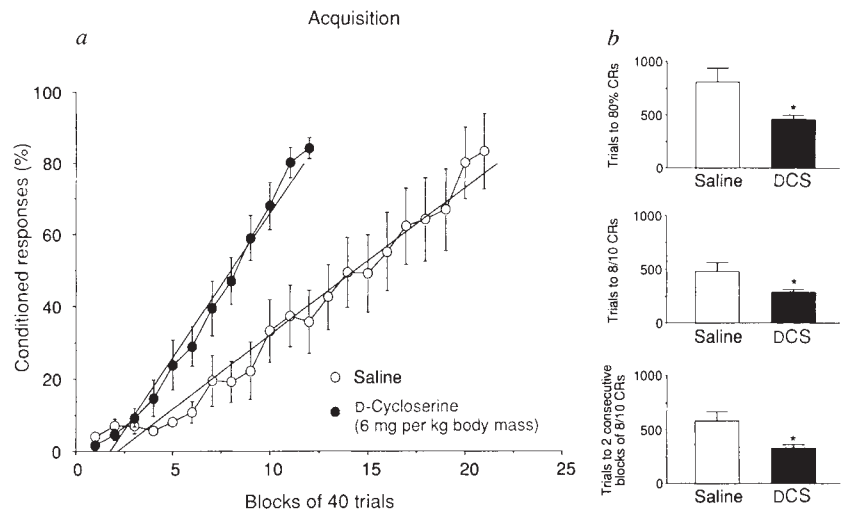


FIG. 3 Daily treatment with D-cycloserine, partial agonist of the glycine coagonist site on the NMDA receptor, also facilitated trace eye-blink conditioning in rabbits ($n=7$ per group). *a*, D-cycloserine significantly accelerated the learning rates of rabbits as compared with saline-treated controls ($P<0.02$). The learning curves shown were normalized (as in Fig. 1) to the mean trials to the criterion level for each group, to allow qualitative comparisons between groups. *b*, D-cycloserine significantly reduced the number of trials required to acquire successfully the trace eye-blink conditioning task ($P<0.01$). Note the similarities in behavioural facilitation obtained with these two chemically dissimilar compounds. Both glycine partial agonists reduced the training required to learn successfully the eye-blink task by nearly half, as compared with controls. Normal adaptive CRs were exhibited by the rapidly learning animals, with nictitating membrane extension beginning shortly before air-puff onset, to protect the cornea from the US. A large number of double blinks by B6B21 and D-cycloserine-treated but not control rabbits were observed within the trace interstimulus period (over 50% of all CRs by the final training session). The first of these double blinks typically occurred shortly after CS onset, the second with a normal response latency (≤ 150 ms before air-puff US onset). Other evidence (E. Akase and J. F. Disterhoft, unpublished data) suggests that these robust double blink responses may be driven by enhanced neural activity in the hippocampus during the trace interval. We have noted a novel short-latency burst of activity by rabbit hippocampal CA1 single-units during or shortly after CS



presentation in trace conditioning, followed by a long-latency burst of firing during the CR that matched earlier reports of behavioural modelling by hippocampal neurons during delay conditioning³⁷. The behavioural responses suggest that hippocampal function during the trace interval is enhanced by both compounds. The horizontal scale differs from Fig. 1, as rabbits without intraventricular cannulae (D-cycloserine studies) learned faster than those with cannulae.

B6B21, D-cycloserine and other glycine-like analogues should be useful in further study of the role of NMDA receptors in learning, as well as the functional deficits associated with ageing³², ischaemia^{2,33} and Alzheimer's dementia³⁴. □

Received 8 May; accepted 27 August 1992.

- Haring, R., Stanton, P., Scheidele, M. & Moskal, J. *J. Neurochem.* **57**, 323-332 (1991).
- Patel, J., Zinkland, W. C., Thompson, C., Keith, R. & Salama, A. *J. Neurochem.* **54**, 849-854 (1990).
- Moyer, J. R., Jr, Deyo, R. A. & Disterhoft, J. F. *Behav. Neurosci.* **104**, 243-252 (1990).
- Monaghan, D. T. & Cotman, C. W. *J. Neurosci.* **5**, 2909-2919 (1985).
- Collingridge, G. L., Kehl, S. J. & McLennan, H. *J. Physiol. (Lond.)* **334**, 33-46 (1983).
- Oliver, M. W., Kessler, M., Larson, J., Schottler, F. & Lynch, G. *Synapse* **5**, 265-270 (1990).
- Thomson, A. M., Walker, V. E. & Flynn, D. M. *Nature* **338**, 422-424 (1989).
- Andersen, P. & Hvalby, Ø. in *Excitatory Amino Acids* (eds Meldrum, B. S., Moroni, F., Simon, R. P. & Woods, J. H.) 463-44 (Raven, New York, 1991).
- Morris, R. G. M., Anderson, E., Lynch, G. S. & Baudry, M. *Nature* **319**, 774-776 (1986).
- Robinson, G. S., Crooks, G. B., Shinkman, P. G. & Gallagher, M. *Psychobiology* **17**, 156-164 (1989).
- Contreras, P. C. *Neuropharmacology* **29**, 291-293 (1990).
- Napier, R., MacRae, M. & Kehoe, E. *J. exp. Psychol., Anim. Behav. Processes* **18**, 182-192 (1992).
- Hood, W. F., Compton, R. P. & Monahan, J. B. *Neurosci. Lett.* **98**, 91-95 (1989).
- Hanngren, H., Hansson, E. & Ullberg, S. *Antibiotics Chemother.* **12**, 46-54 (1961).
- Monahan, J. B., Corpus, V. M., Hood, W. F., Thomas, J. W. & Compton, R. P. *J. Neurochem.* **53**, 370-375 (1989).
- Watson, G. B., Bolanowski, M. A., Baganoff, M. P., Deppeler, C. L. & Lanthorn, T. H. *Brain Res.* **510**, 158-160 (1990).
- Golub, E. S. *Immunology: A Synthesis* (Sinauer, Sunderland, MA, 1987).
- Vartanian, M. G. & Taylor, C. P. *Neurosci. Lett.* **133**, 109-112 (1991).
- Monahan, J. B., Handelman, G. E., Hood, W. F. & Cordi, A. A. *Pharmac. biochem. Behav.* **34**, 649-653 (1989).
- Schwartz, B. L., Hashtroudi, S., Herting, R. L., Handerson, H. & Deutsch, S. I. *Neurology* **41**, 1341-1343 (1991).
- Thiels, E., Weisz, D. J. & Berger, T. W. *Neuroscience* **46**, 501-509 (1992).
- Disterhoft, J. F., Coulter, D. A. & Alkon, D. L. *Proc. natn. Acad. Sci. U.S.A.* **83**, 2733-2737 (1986).
- de Jonge, M. C., Black, J., Deyo, R. A. & Disterhoft, J. F. *Expl. Brain Res.* **80**, 456-462 (1990).
- Solomon, P. R., van der Schaaf, E. V., Thompson, R. F. & Weisz, D. J. *Behav. Neurosci.* **100**, 729-744 (1986).
- Gaffan, D. & Gaffan, E. A. *Brain* **114**, 2611-2618 (1991).
- Squire, L. *Memory and Brain* (Oxford, Univ. Press, London, 1987).
- Dedrick, R. L. & Flessner, M. F. in *Immunity to Cancer II* (ed. Mitchell, M. S.) 429-438 (Liss, New York, 1989).
- Winter, G. & Milstein, C. *Nature* **349**, 293-299 (1991).
- Saragovi, H. U. et al. *Science* **253**, 792-795 (1991).
- Woodruff-Pak, D. S. & Thompson, R. F. *Psychol. Aging* **3**, 219-229 (1988).
- Solomon, P. R., Levine, E., Bein, T. & Pendlebury, W. W. *Neurobiol. Aging* **12**, 283-287 (1991).
- Miyoshi, R., Kito, S., Doudou, N. & Nomoto, T. *Synapse* **6**, 338-343 (1990).
- Foster, A. C., Willis, C. L. & Tridgett, R. *Eur. J. Neurosci.* **2**, 270-277 (1990).
- Procter, A. W., Stirling, J. M., Stratmann, G. C., Cross, A. J. & Bowen, D. M. *Neurosci. Lett.* **101**, 62-66 (1989).
- Moskal, J. R. & Schaffner, A. E. *J. Neurosci.* **6**, 2045-2063 (1986).
- Deyo, R. A., Straube, K. & Disterhoft, J. F. *Science* **243**, 809-811 (1989).
- Berger, T. W., Rinaldi, P. C., Weisz, D. J. & Thompson, R. F. *J. Neurophysiol.* **50**, 1197-1219 (1983).

ACKNOWLEDGEMENTS. This project was supported by grants from the National Institute of Aging, the National Institute of Drug Abuse, and the Office of Naval Research. We thank T. Lanthorn for comments and S. Conroy, M. McCarthy, and S. Loughran for technical assistance.

Deletion polymorphism in the gene for angiotensin-converting enzyme is a potent risk factor for myocardial infarction

François Cambien^{*†}, Odette Poirier^{*}, Laure Lecerf^{*}, Alun Evans[‡], Jean-Pierre Cambou[§], Dominique Arveiler^{||}, Gerald Luc[¶], Jean-Marie Bard[#], Lucienne Bara^{**}, Sylvain Ricard^{*}, Laurence Tiret^{††}, Philippe Amouyel[¶], François Alhenc-Gelas^{‡‡} & Florent Soubrier^{‡‡}

^{*} INSERM SC7, 17 rue du Fer à Moulin, 75005 Paris, France

[†] INSERM U258, Hôpital Broussais, 96 rue Didot, 75674 Paris, France

[‡] MONICA Project, The Queen's University, Belfast BT12 6BJ, UK

[§] MONICA Project, INSERM U326, CHU Purpan, 31059 Toulouse, France

^{||} MONICA Project, 11 rue Humann, 67085 Strasbourg, France

[¶] MONICA Project, Institut Pasteur, 59019 Lille, France

[#] SERLIA-INSERM U325, Institut Pasteur, 59019 Lille, France

^{**} Laboratoire de Thrombose Expérimentale,

15 rue de l'Ecole de Médecine, 75005 Paris, France

^{††} INSERM U367, 17 rue du Fer à Moulin, 75005 Paris, France

^{‡‡} INSERM U36, 3 rue d'Ulm, 75005 Paris, France

FACTORS involved in the pathogenesis of atherosclerosis, thrombosis and vasoconstriction^{1,2} contribute to the development of coronary heart disease. In a study comparing patients after myocardial infarction with controls, we have explored a possible association between coronary heart disease and a variation found in the gene encoding angiotensin-converting enzyme (ACE). The polymorphism ACE/ID is strongly associated with the level of circulating enzyme³. This enzyme plays a key role in the production of angiotensin II and in the catabolism of bradykinin, two peptides involved in the modulation of vascular tone and in the proliferation of smooth muscle cells. Here we report that the DD genotype, which is associated with higher levels of circulating ACE than the ID and II genotypes, is significantly more frequent in patients

TABLE 1 Distribution of ACE genotypes in cases and controls in the different populations of the ECTIM study

ACE genotype	Belfast		Lille		Strasbourg		Toulouse		All subjects	
	Cases	Controls	Cases	Controls	Cases	Controls	Cases	Controls	Cases	Controls
Number of subjects										
DD	51	43	27	43	67	55	52	59	197	200
ID	111	95	24	73	107	98	67	124	309	390
II	39	42	7	32	30	41	28	28*	104	143
Odds ratios†	1.1		2.1		1.2		1.4		1.34‡	
D/I	0.53/0.47	0.50/0.50	0.67/0.33	0.54/0.46	0.59/0.41	0.54/0.46	0.58/0.42	0.57/0.43		

Statistical analysis of association between myocardial infarction (MI) and ACE polymorphism (logistic regression analysis):

ACE genotype (DD more frequent in cases than in controls): $P=0.007$.

Population (different relative frequency of cases and controls in the 4 populations): $P<0.0001$.

Heterogeneity of ACE genotype effect across populations: $P=0.4$.

The populations participating in the ECTIM study have been described and their representativity analysed in detail²². Briefly, cases were recruited from the MONICA registers²³ in Belfast (Northern Ireland), Lille (northern France), Strasbourg (eastern France) and Toulouse (south western France). Patients aged 25 to 64 who survived a myocardial infarction (MI) according to the MONICA diagnostic category I over the study period (December 1988–June 1991) were eligible. Patients were recruited in hospitals and blood samples taken at least 3 months and at most 9 months after their MI. The samples were obtained from controls, from electoral rolls in Strasbourg, Lille and Toulouse and from the lists of General Practitioners held by the Central Services Agency in Belfast. Informed consent was obtained from the subjects. To increase the ethnic homogeneity within the regions, families of patients and controls had to have been resident in the region for at least two generations and all their grandparents had to have been born in Europe; all participants were caucasians. The numbers of patients with MI and controls included in this analysis of the ACE gene polymorphism were, respectively, 201 and 180 from Belfast, 58 and 148 from Lille (the relatively small number of cases in this centre was due to unexpected difficulties of recruitment), 204 and 194 from Strasbourg, and 147 and 211 from Toulouse. The mean age of cases and controls was similar: 53.9 (s.d. 8.8) and 53.1 (s.d. 8.5), respectively, and the mean cholesterol level in the control groups was identical to those obtained in the random samples of the first MONICA survey²². The distribution of ACE genotypes was independent of age and social class.

* Significant departure from Hardy-Weinberg equilibrium: $P<0.005$.

† The odds ratio is a measure of relative risk of MI between subjects with the DD genotype and subjects with the ID or II genotype.

‡ Odds ratio adjusted on population.

with myocardial infarction ($n=610$) than in controls ($n=733$) ($P=0.007$), especially among subjects with low body-mass index and low plasma levels of ApoB ($P<0.0001$). The ACE/ID polymorphism seems to be a potent risk factor of coronary heart disease in subjects formerly considered to be at low risk according to common criteria.

The level of plasma angiotensin-converting enzyme is stable in an individual but is highly variable between individuals⁴. About 50% of this inter-individual variability of plasma ACE concentration is determined by a major gene effect⁵. After the ACE gene was cloned⁶ it was shown that this major gene effect is associated with an insertion (I)/deletion (D) polymorphism situated in intron 16 of the ACE gene (ACE/ID polymorphism)⁷; the mean plasma ACE level in DD subjects was about

twice that of II subjects, with ID subjects having intermediate levels³.

As the ACE gene is a candidate gene for cardiovascular disease, any possible association between the ACE/ID polymorphism and myocardial infarction was investigated in the ECTIM (Etude Cas-Témoin de l'Infarctus du Myocarde) study, a multicentre study designed to identify variants of candidate genes predisposing to myocardial infarction. This study included male patients and appropriate controls recruited in different European populations at highly contrasting risk for coronary heart disease. The ACE/ID polymorphism was analysed by DNA amplification⁸.

The distributions of the ACE genotypes (Table 1) in cases and controls in the different populations were in Hardy-Weinberg equilibrium, except in the control group of Toulouse where, for unknown reasons, an excess of heterozygotes was observed. In the whole population, the genotype characterized by two short alleles (genotype DD) was associated with an excess of cases with myocardial infarction compared with the two other genotypes (ID and II) ($P=0.007$). This association was not significantly heterogeneous across populations and the frequency of the two alleles was similar in the different control groups.

In view of the association between the ACE/ID polymorphism and myocardial infarction, different variables potentially related to the disease were compared between ACE/ID genotypes. As shown in Table 2, no difference could be detected between genotypes for plasma lipid variables, fibrinogen, body mass index (BMI), cigarette consumption or blood pressure. Moreover, the prevalence of hypertension, as defined by a diastolic blood pressure higher than 100 mm Hg or by antihypertensive treatment, was similar in the three genotypes: 22.0%, 17.7% and 20.3% in genotypes DD, ID and II respectively (nonsignificant differences).

As the influence of the DD genotype seemed to be independent of classical risk factors, we considered that it might be particularly important in myocardial infarction patients considered to be at low risk according to the usual criteria. Hyperlipidaemia and obesity are important contributors to the coronary heart disease epidemic in western countries, so a

TABLE 2 Comparison of several variables between ACE/ID genotypes in the control group of the ECTIM study

	Numbers	ACE genotype		
		DD	ID	II
		Means (s.e.m.)	adjusted on population and age	
Total cholesterol (mg dl ⁻¹)	731	230.9 (3.2)	231.9 (3.2)	230.9 (3.8)
ApoA1 (mg dl ⁻¹)	731	147.8 (1.9)	148.0 (1.4)	145.5 (2.2)
ApoB (mg dl ⁻¹)	731	130.3 (2.7)	129.4 (1.9)	128.3 (3.2)
Lp(a) (mg dl ⁻¹)	729	10.6 (1.2)	10.2 (0.9)	9.8 (1.4)
Fibrinogen (mg dl ⁻¹)	663	307.2 (5.9)	313.6 (4.2)	313.9 (6.8)
Systolic blood pressure (mm Hg)	733	131.7 (1.3)	132.3 (1.0)	134.2 (1.6)
Diastolic blood pressure (mm Hg)	733	81.1 (0.9)	81.5 (0.7)	83.0 (1.1)
Body mass index (kg m ⁻²)	733	26.2 (0.3)	26.3 (0.2)	26.8 (0.2)
Cigarette consumption (per day)	733	5.6 (0.7)	5.6 (0.5)	5.7 (0.9)

This analysis was performed in the control groups to avoid the potential bias that might result from lifestyle changes (particularly treatment) occurring after myocardial infarction. No statistically significant difference was detected between ACE genotypes. The techniques used in the ECTIM study to measure lipid variables have been described²². Fibrinogen was measured by the thrombin time method. Blood pressure is the mean of two measurements made at 5-min intervals with a random zero sphygmomanometer on subjects in the sitting position at rest.

TABLE 3 Distribution of ACE genotypes in cases and controls in the different populations according to risk status

	Belfast		Lille		Strasbourg		Toulouse		All subjects	
	Cases	Controls	Cases	Controls	Cases	Controls	Cases	Controls	Cases	Controls
ACE genotype					Low-risk group*					
DD	13	12	5	11	10	7	10	16	38	46
ID + II	15	40	1	28	14	28	11	47	41	143
Odds ratios†	2.9 (1.0–8.7)		12.7‡		2.9 (0.8–10.7)		2.7 (0.8–8.4)		3.2 (1.7–5.9)§	
					High-risk group					
DD	38	31	22	32	57	48	42	43	159	154
ID + II	135	97	30	77	123	111	84	105	372	390
Odds ratios	0.9 (0.5–1.6)		1.8 (0.8–3.7)		1.1 (0.7–1.7)		1.2 (0.7–2.1)		1.1 (0.9–1.5)	

Statistical analysis of association between myocardial infarction and ACE polymorphism (multiple logistic regression analysis):

ACE genotype (DD more frequent in cases than in controls): $P < 0.0001$ ($P < 0.0005$)¶.

Population (different relative frequency of cases and controls in the 4 populations): $P < 0.0001$ ($P < 0.05$).

Risk group (high-risk status more frequent in cases than in controls): $P < 0.0001$ ($P < 0.03$).

Heterogeneity of ACE genotype effect across populations: $P = 0.3$ ($P = 0.4$).

Heterogeneity of risk group effect across populations: $P = 0.4$ ($P = 0.3$).

Heterogeneity of ACE genotype effect according to risk group (DD genotype more strongly associated with MI in the low-risk group): $P < 0.001$ ($P < 0.02$).

* Subjects with BMI $< 26 \text{ kg m}^{-2}$, ApoB level $< 125 \text{ mg dl}^{-1}$ and not treated with hypolipidaemic drugs.

† 95% confidence interval in brackets.

‡ Fisher exact test, $P = 0.03$.

§ Odds-ratio adjusted on population.

|| Subjects who are not in the 'low risk' group.

¶ After adjustment for age, cigarette consumption, systolic blood pressure, social class and plasma ApoA1 and Lp(a) levels.

low-risk group was defined according to plasma ApoB level and BMI. This low-risk group included subjects with a plasma ApoB level lower than 125 mg dl^{-1} (roughly the median level of ApoB in controls), those not undergoing treatment with hypolipidaemic drugs and those with a BMI $< 26 \text{ kg m}^{-2}$ (the roughly median BMI in controls). In this group (Table 3), the association between the ACE/ID polymorphism and myocardial infarction was highly significant and was homogeneous across the populations; the overall odds ratio (an estimate of relative risk), adjusted for population, was 3.2 ($P < 0.0001$). Conversely, in the high-risk group, which included all subjects who were not in the low-risk group already defined (Table 3), the ACE/ID polymorphism was not related to myocardial infarction. In the whole population, after adjustment for age, cigarette consumption, plasma ApoA1, plasma Lp(a), blood pressure and social class, the DD genotype and its interaction with the risk status were still highly significantly associated with myocardial infarction ($P < 0.0005$ and $P < 0.02$, respectively). The ACE/ID polymorphism is thus an independent risk factor for myocardial infarction in low-risk individuals.

We defined our low-risk group by reference to plasma ApoB level, hypolipidaemic treatment and BMI; the results were very similar (but slightly less significant) when total cholesterol or cholesterol in low-density lipoproteins (LDL) was used instead of ApoB. In contrast, the association between the DD genotype and myocardial infarction was similar in the groups defined according to the median of cigarette consumption, ApoA1, Lp(a) or fibrinogen. Furthermore, as many of the patients received potent drugs likely to affect blood pressure, we were unable to define a normotensive low-risk group.

It is unlikely that population heterogeneity could explain our results: in each population, cases and controls were carefully selected to minimize ethnic heterogeneity. All subjects participating in our study were caucasians; we have studied other samples from general populations of caucasian origin and found that the frequency of the ACE/ID polymorphism was fairly similar to that in the ECTIM study. The fact that the association was homogeneous across populations also provides a strong argument against a possible confounding effect of population heterogeneity.

The frequency of the ACE/DD genotype in the general population was 0.27. If we assume that the relative risk is well approximated by the odds ratio, the percentage of cases that

could be attributable to this genotype was 8% in the whole population and 35% in the low-risk group. The values of these estimates indicate a strong impact of the ACE/ID polymorphism on myocardial infarction in the population.

ACE exists predominantly as an ectoenzyme of vascular endothelial cells and plays a key part in the renin-angiotensin and kallikrein-kinin systems by activating angiotensin I into angiotensin II and inactivating bradykinin. These two peptide hormones have opposite effects on vascular tone and on smooth muscle cell proliferation^{9–15}, and as neointimal proliferation and vasospasm may be involved in the pathogenesis of coronary heart disease^{1,2}, the most likely mechanism by which the ACE polymorphism affects the risk of myocardial infarction is by modulating the level of these peptides in the coronary arteries. This hypothesis is compatible with earlier results¹⁶ and with a study¹⁷ of hypertensive patients with high renin profiles, a condition likely to be associated with increased angiotensin II, who were found to be at higher risk of coronary heart disease than those with low renin profiles. But we cannot exclude an effect of ACE on the metabolism of other peptides in view of its wide substrate specificity¹⁸, nor can we exclude the possibility that a variant of another gene, closely linked to the ACE gene, causes the observed association.

The fact that the ACE polymorphism is unrelated to blood pressure and hypertension agrees with results obtained from hypertensive sibs¹⁹. An association between the ACE/ID polymorphism and blood pressure might have been expected from its anticipated effect on vascular tone and architecture. But the deleterious effect of the DD genotype might be restricted to certain vascular territories like the coronary circulation²⁰ or might develop preferentially in atheromatous arteries. Also, regulatory mechanisms such as increased natriuresis may counteract the effect of ACE on peripheral resistance.

Results from a family study combining linkage and segregation analyses have indicated that the ACE/ID polymorphism is unlikely to be the locus directly affecting the variability of plasma ACE, but that it could be a marker in strong linkage disequilibrium with this locus⁷. We have inferred that the sensitivity and specificity of the ACE/D allele as a marker of the postulated 'susceptibility' allele were 1 and 0.77, respectively. So the ACE/D allele seems to be a very good marker for the susceptibility allele in that population, which may not be the case in non-caucasian populations. Characterization of the

molecular basis of the susceptibility genotype should help explain why the risk of myocardial infarct is increased in *DD* individuals and improve the specificity of risk assessment. The plasma ACE level could turn out to be more useful for risk assessment than identifying the ACE genotype.

In conclusion, the *ACE/DD* genotype is a new potent risk factor for myocardial infarction in middle-aged men, par-

ticularly in subjects normally considered to be at low risk. The genetic nature of this risk factor does not preclude the possibility of prevention because ACE blockers are available, but their effectiveness and safety for preventing myocardial infarction remain to be established²¹. As with any newly proposed risk factor, independent confirmation of our results will be necessary. □

Received 8 June; accepted 10 September 1992.

1. Maseri, A. *et al.* *New Engl. J. Med.* **299**, 1271-1277 (1978).
2. Fuster, V., Badimon, L., Badimon, J. J. & Chesebro, J. H. *New Engl. J. Med.* **326**, 310-318 (1992).
3. Rigat, B. *et al.* *J. Clin. Invest.* **86**, 1343-1346 (1990).
4. Alhenc-Gelas, F., Richard, J., Courbon, D., Warnet, J. M. & Corvol, P. *J. lab. clin. Med.* **117**, 33-39 (1991).
5. Cambien, F. *et al.* *Am. J. hum. Genet.* **43**, 774-780 (1988).
6. Soubrier, F. *et al.* *Proc. natn. Acad. Sci. U.S.A.* **85**, 9386-9390 (1988).
7. Tiret, L. *et al.* *Am. J. hum. Genet.* **51**, 197-205 (1992).
8. Rigat, B., Hubert, C., Corvol, P. & Soubrier, F. *Nucleic Acids Res.* **20**, 1433 (1992).
9. Daemen, M. J. A. P., Lombardi, D. M., Borman, F. T. & Schwartz, S. M. *Circ. Res.* **68**, 450-456 (1991).
10. Powell, J. S. *et al.* *Science* **245**, 186-188 (1989).
11. Berk, B. C., Vekshtein, V., Gordon, H. M. & Tsuda, T. *Hypertension* **13**, 305-314 (1989).
12. Kato, H. *et al.* *J. Hypertens.* **9**, 17-22 (1991).
13. Pelc, L. R., Gross, G. J. & Warltier, D. C. *Circulation* **83**, 2048-2056 (1991).
14. Wiemer, G., Schölkens, B. A., Becker, R. H. A. & Busse, R. *Hypertension* **18**, 558-563 (1991).

15. Farhy, R. D., Ho, K. L., Carretero, O. A. & Scicli, A. G. *Biochem. Biophys. Res. Commun.* **182**, 283-288 (1992).
16. Brunner, H. R. *et al.* *New Engl. J. Med.* **286**, 441-449 (1972).
17. Alderman, M. H. *et al.* *New Engl. J. Med.* **324**, 1098-1104 (1991).
18. Erdős, E. G. *Hypertension* **16**, 363-370 (1990).
19. Jeunemaitre, K., Lifton, R. P., Hunt, S. C., Williams, R. R. & Lalouel, J. M. *Nature Genet.* **1**, 72-75 (1992).
20. Magrini, F. *et al.* *Circulation* **75** (suppl.), 168-174 (1987).
21. McMurray, J., MacLenachan, J. & Dargie, H. J. *J. Hypertens.* **9**, 393-397 (1991).
22. Parra, H. J. *et al.* *Arterioscl. Thromb.* **12**, 701-707 (1992).
23. WHO MONICA Project Principal Investigators. *J. Clin. Epidemiol.* **41**, 105-114 (1988).

ACKNOWLEDGEMENTS. We thank C. Souriau and A. Micael for technical work, and P. Ducimetière, P. Corvol and J. Ménard for comments on the manuscript. The ECTIM study is supported by grants from the Squibb laboratory, Sandoz laboratory, the Mutuelle Générale de l'Éducation Nationale, the British Heart Foundation, and by the Institut National de la Santé et de la Recherche Médicale and the Institut Pasteur, Lille.

Oxidative stress and heat shock induce a human gene encoding a protein-tyrosine phosphatase

Stephen M. Keyse & Elizabeth A. Emslie

ICRF Molecular Pharmacology Unit, University Department of Biochemistry, Hugh Robson Building, George Square, Edinburgh EH8 9XD, UK

REACTIVE oxygen species have been implicated both in the ageing process and in degenerative diseases, including arthritis and cancer^{1,2}. Bacteria adapt to the lethal effects of oxidants such as hydrogen peroxide by inducing the expression of protective stress genes^{3,4}. Analogous responses have been identified in human cells. For example, haem oxygenase is a major stress protein in human cells treated with oxidants⁵, and reactive oxygen intermediates activate NF- κ B, a transcriptional regulator of genes involved in inflammatory and acute-phase responses⁶. We report here the isolation and characterization of a novel complementary DNA (CL100) corresponding to a messenger RNA that is highly inducible by oxidative stress and heat shock in human skin cells. The cDNA contains an open reading frame specifying a protein of *M_r* 39.3K with the structural features of a non-receptor-type protein-tyrosine phosphatase⁷ and which has significant amino-acid sequence similarity to a Tyr/Ser-protein phosphatase encoded by the late gene *H1* of vaccinia virus⁸. The purified protein encoded by the CL100 open reading frame expressed in bacteria has intrinsic phosphatase activity. Given the relationship between the levels of protein-tyrosine phosphorylation, receptor activity, cellular proliferation and cell-cycle control, the induction of this gene may play an important regulatory role in the human cellular response to environmental stress.

To isolate cDNAs corresponding to genes induced by oxidative stress, a cDNA library was prepared from human skin fibroblasts treated with 100 μ M hydrogen peroxide. This library was differentially screened with a probe of total cDNA from peroxide-treated cells and a probe from untreated cells. One clone (designated CL100) reproducibly showed a much higher hybridization signal with the induced probe relative to control and was characterized further. Northern blot analysis shows that the cDNA insert from CL100 hybridizes with a single, highly inducible mRNA of about 2.4 kilobases (kb) (Fig. 1a). This mRNA accumulates rapidly after treatment with hydrogen

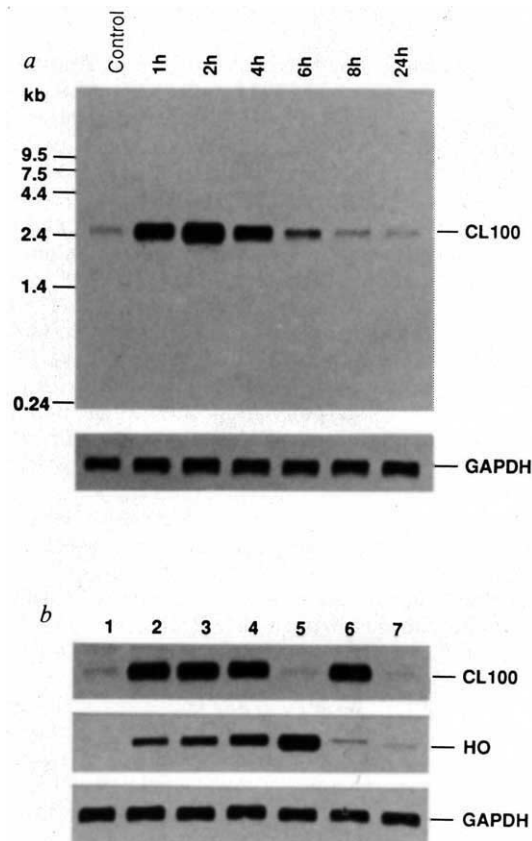


FIG. 1 a, Time course of CL100 mRNA accumulation in human skin fibroblasts (strain EK4)²⁴ following treatment with 200 μ M hydrogen peroxide. b, The accumulation of CL100 mRNA in EK4 fibroblasts 2 h after various chemical and physical treatments compared with the levels of mRNA for the stress-inducible haem oxygenase (HO)⁵ gene. Lanes 1 and 7, control; lane 2, 200 μ M hydrogen peroxide; lane 3, 500 μ M menadione; lane 4, 8×10^5 J m⁻² of ultraviolet A (320-380 nm) radiation; lane 5, 50 μ M sodium arsenite; lane 6, heat shock 45°C for 5 min. Both blots were reprobbed with the *Pst*I fragment of the glyceraldehyde phosphate dehydrogenase gene (GAPDH)²⁵ as a loading control.

METHODS. Cells were treated with chemicals for 30 min at 37°C or as stated and then incubated for the indicated times before isolation of total RNA²⁶. Northern blot analysis was done using standard techniques²⁷.

TABLE 1 Induction of CL100 mRNA by DNA-damaging agents

Treatment	Level of induction
H ₂ O ₂ (200 μ M)	23.4
Menadione (500 μ M)	24.2
UVA (8×10^5 J m ⁻²)	18.2
UVC (10 J m ⁻²)	1.6
MNNG (10 μ M)	1.1
<i>cis</i> -DDP (50 μ M)	1.6
X-rays (2.0 G)	1.4

Levels of CL100 mRNA were determined 2 h after the above treatments by northern blotting, exactly as described in Fig. 1. Hybridization signals were quantified by densitometric scanning of these blots and levels of induction are expressed as fold increase over the signals in untreated controls. The figures given are readings from single experiments, typical of 2 or more independent determinations. H₂O₂, Hydrogen peroxide; UVC, far ultraviolet radiation (254 nm); UVA, near ultraviolet radiation (320–380 nm); MNNG, *N*-methyl-*N*'nitro-*N*-nitrosoguanidine; *cis*-DDP, *cis*-diaminedichloro-platinum(II).

peroxide, reaching levels up to 40-fold greater than those seen in untreated cells within 2 h. Elevated levels of the transcript persist for at least 6 h before returning to normal.

Cells were treated with a range of chemical and physical insults before analysis of mRNA (Fig. 1b) and the pattern of induction of CL100 mRNA was compared with that for haem oxygenase, a gene transcriptionally induced in human cells by oxidant stress⁹. In addition to hydrogen peroxide, CL100 mRNA is strongly induced by two other oxidants: menadione, a superoxide-generating drug; and radical-inducing ultraviolet A (320–380 nm) radiation. CL100 mRNA levels are also increased in response to heat shock. Although both CL100 and haem oxygenase mRNAs are induced by oxidants they are not coordinately induced by other stress treatments.

A number of genes have been identified as inducible in mammalian cells by agents that cause DNA damage and growth arrest. These include DNA polymerase- β ¹⁰, *c-fos*¹¹ and *c-jun*¹². Because oxidants damage DNA and are also mutagenic^{13,14} it was of interest to determine whether DNA damage might be causing induction of the CL100 gene. Cells were treated with different DNA-damaging agents and assayed for induction of CL100 mRNA. As several well characterized DNA-damaging treatments led to no induction (Table 1), we conclude that damage to DNA following treatment with oxidants is not the trigger for the induction of this gene.

The nucleotide sequence of the CL100 cDNA was determined (Fig. 2a). The insert is 2 kb long excluding the poly(A)⁺ tail. The first ATG in from the 5' end is located at nucleotide 234 in a sequence context favourable to translation initiation¹⁵. A con-

tinuous open reading frame extends to nucleotide 1,334 and specifies a polypeptide of 367 amino acids with a predicted molecular mass of 39,302. By using the MOTIFS search for protein patterns defined in the PROSITE dictionary¹⁶, we have found that the carboxy terminal portion of the CL100 polypeptide contains a single copy of the highly conserved active-site sequence of protein-tyrosine phosphatases (PTPases) Ile/Val-His-Cys-X-Ala-Gly-X-X-arg (amino acids 256–264)⁷. The putative CL100 protein thus has the structural features of an expanding family of non-receptor-type protein-tyrosine phosphatases which include the *cdc25* M-phase inducer¹⁷, placental PTPase1B¹⁸ and T-cell PTPase¹⁹.

To determine whether the CL100 protein possesses phosphatase activity, the open reading frame (nucleotides 234–1,421) was subcloned into the bacterial expression vector pET-15b (Novagen) and recombinant protein was isolated in a highly purified form (Fig. 3a). This protein rapidly hydrolyses *p*-nitrophenyl phosphate (pNPP), a chromogenic molecule that is structurally related to phosphotyrosine (Fig. 3b). The dephosphorylation of pNPP by the CL100 protein displays many of the properties typical for the known protein-tyrosine phosphatases. Sodium vanadate (a potent inhibitor of protein-tyrosine phosphatases) abolishes the enzyme activity (Fig. 3b). Zinc ions also strongly repress the CL100 phosphatase. Activity is dependent on the presence of reducing agent and shows no requirement for divalent cations (data not shown).

A search of the EMBL/GenBank and SwissProt databases using the FASTA program²⁰ reveals that the CL100 nucleotide sequence is 86.6% identical to a recently characterized mouse cDNA (3CH134) for a growth factor-inducible immediate early gene²¹ and has significant amino-acid sequence similarity to the Ser/Tyr-phosphatase encoded by the late gene *H1* of vaccinia virus⁸. The mouse cDNA encodes a protein of 367 amino acids which is 96.5% identical to the CL100 polypeptide and is clearly the mouse homologue of our human gene. The presence of the PTPase active-site sequence and similarity to the vaccinia protein has also been noted²¹.

The relationship between the amino-acid sequences of CL100, 3CH134 and other non-receptor protein-tyrosine phosphatases is shown in Fig. 2b. The region of similarity between the CL100 protein and the vaccinia H1 sequence extends from residue 182 through the active-site sequence, ending at residue 274. Although over this region there is significant similarity between the vaccinia protein and other human non-receptor PTPases (placental and T cell)⁸ these residues (with the exception of the conserved phenylalanine at position 243) are not present in the CL100 sequence. The significance of the sequence similarity between our gene and the vaccinia protein is as yet unclear. Unlike other PTPases the vaccinia protein is able to dephosphorylate proteins

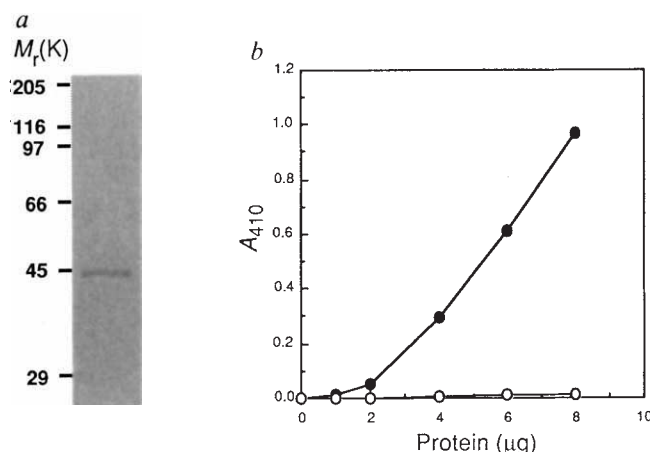


FIG. 3 a, SDS-PAGE of purified His-Tag CL100 fusion protein. M_r s of standards (K) are shown on the left side of the gel. b, The CL100 protein has intrinsic phosphatase activity. The purified protein was assayed at the indicated concentrations for its ability to hydrolyse *p*-nitrophenol phosphate (pNPP) either in the absence (●) or presence (○) of 200 μ M sodium vanadate. METHODS. The CL100 cDNA was used as a template in a polymerase chain reaction to generate an *Nde*I-*Xho*I fragment containing the CL100 open reading frame (nucleotides 234–1,421). This was then inserted into *Nde*I-*Xho*I digested pET-15b (Novagen) vector DNA and the reading frame verified by DNA sequence analysis. This plasmid was used to express CL100 as a His-Tag fusion protein which was purified in a denatured form as described³⁰. The purified protein was then refolded by gradual removal of the denaturant by dialysis³¹ and analysed by SDS-PAGE. Hydrolysis of pNPP was in a reaction volume of 200 μ l containing 50 mM imidazole, pH 7.5, 0.1% β -mercaptoethanol, 20 mM pNPP and enzyme at 30 °C for 30 min. The reaction was stopped by addition of 800 μ l of 0.25 M NaOH and the absorbance at 410 nm (A_{410}) was measured.

modified both at serine and at tyrosine residues, but the structural basis of this altered specificity and its biological significance are unknown⁸. We conclude that the CL100 gene and its mouse homologue represent a new subclass of mammalian PTPase in that they show no significant similarity to any human PTPase yet discovered other than the presence of the highly conserved active-site sequence.

Tyrosine phosphorylation and dephosphorylation are central

to the process of growth regulation, differentiation and oncogenesis^{22,23}. We speculate that the PTPase encoded by the CL100 gene may play an important regulatory role in the cellular response to adverse environmental challenges, including oxidants and hyperthermia. Further study of the biological activity of the CL100 protein and the identification of its physiological substrates should yield vital information as to the nature of the signal transduction pathways involved. □

Received 18 May; accepted 24 August 1992.

- Ames, B. N. in *Genetic Toxicology of Environmental Chemicals* (eds Black, L. & Nathan, K.) 11–26 (Liss, New York, 1985).
- Halliwel, B. & Gutteridge, J. M. C. *Meth. Enzym.* **186**, 1–85 (1990).
- Christman, M. F., Morgan, R. W., Jacobson, F. S. & Ames, B. N. *Cell* **41**, 753–762 (1985).
- Greenberg, J. T., Monach, P. A., Chou, J. H., Josephy, P. D. & Demple, B. *Proc. natn. Acad. Sci. U.S.A.* **87**, 6181–6185 (1990).
- Keyse, S. M. & Tyrrell, R. M. *Proc. natn. Acad. Sci. U.S.A.* **86**, 99–103 (1989).
- Schreck, R., Rieber, P. & Baeuerle, P. A. *EMBO J.* **10**, 2247–2258 (1991).
- Fischer, E. H., Charbonneau, H. & Tonks, N. K. *Science* **253**, 401–406 (1991).
- Guan, K., Broyles, S. S. & Dixon, J. E. *Nature* **350**, 359–362 (1991).
- Keyse, S. M., Applegate, L. A., Tromvoukis, Y. & Tyrrell, R. M. *Molec. cell. Biol.* **10**, 4967–4969 (1990).
- Fornace, A. J., Zmudka, B. Z., Hollander, M. C. & Wilson, S. H. *Molec. cell. Biol.* **9**, 851–853 (1989).
- Boscher, M., Rahmsdorf, H. J., Litfin, M. & Herrlich, P. *Oncogene* **3**, 301–311 (1988).
- Sherman, M. L., Datta, R., Hallahan, D. E., Weischelbaum, R. R. & Kufe, D. W. *Proc. natn. Acad. Sci. U.S.A.* **87**, 5663–5666 (1990).
- Hutchinson, F. *Prog. Nucleic Acids Res.* **32**, 115–154 (1985).
- Ziegler-Skylakakis, K. & Andrae, U. *Mutat. Res.* **192**, 65–67 (1987).
- Kozak, M. *Cell* **44**, 283–292 (1986).
- Bairoch, A. *Prosite: A Dictionary of Protein Sites and Patterns* 8th release (University of Geneva, 1991).
- Millar, B. A. & Russell, P. *Cell* **68**, 407–410 (1992).

- Chernoff, J., Schievella, A. R., Jost, C. A., Erikson, R. L. & Neel, B. G. *Proc. natn. Acad. Sci. U.S.A.* **87**, 2735–2739 (1990).
- Cool, D. E. et al. *Proc. natn. Acad. Sci. U.S.A.* **86**, 5257–5261 (1989).
- Pearson, W. R. & Lippman, D. J. *Proc. natn. Acad. Sci. U.S.A.* **85**, 2444–2448 (1988).
- Charles, C. H., Abler, A. S. & Lau, L. F. *Oncogene* **7**, 187–190 (1992).
- Hunter, T. A. & Cooper, J. A. *Rev. Biochem.* **54**, 897–930 (1985).
- Cantley, L. C. et al. *Cell* **64**, 281–302 (1991).
- Tyrrell, R. M. & Pidoux, M. *Cancer Res.* **44**, 2665–2669 (1986).
- Piechaczyk, M. et al. *Nucleic Acids Res.* **12**, 6951–6963 (1984).
- Chomczynski, P. & Sacchi, N. *Analyt. Biochem.* **162**, 156–159 (1987).
- Maniatis, T., Fritsch, E. F. & Sambrook, J. *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Laboratory Press, New York, 1982).
- Shaw, G. & Kamen, R. *Cell* **46**, 659–667 (1986).
- Leonard, D. G. B., Ziff, E. B. & Greene, L. A. *Molec. cell. Biol.* **7**, 3156–3167 (1987).
- Hochuli, E., Bannwarth, W., Dobeli, H., Gentz, R. & Stuber, D. *Biotechnology* **6**, 1321–1325 (1988).
- Clackson, L. A., Ahn, B., Koo, H. S. & Grossman, L. *J. biol. Chem.* **266**, 11380–11387 (1991).

ACKNOWLEDGEMENTS. We thank R. M. Tyrrell for RNA samples, H. Lohrer for assistance with X-ray experiments, J. E. Freeman for photography, I. J. Goldsmith for oligonucleotides, T. Lindahl and R. D. Wood for discussion and encouragement, and C. R. Wolf for reading the manuscript. This work was supported by the ICRF. The nucleotide sequence data will appear in the EMBL, GenBank and DDBJ nucleotide sequence databases under accession number X68277.

A functional barrier to movement of lipids in polarized neurons

Toshihide Kobayashi*, Brian Storrie†, Kai Simons & Carlos G. Dotti‡

Cell Biology Program, European Molecular Biology Laboratory, Postfach 10.2209, D-6900 Heidelberg, Germany

† Department of Biochemistry, Virginia Polytechnic Institute and State University, Blacksburg, Virginia 24061-0308, USA

In polarized neurons, axons and dendrites perform different functions, which are reflected in their different molecular organization. Studies on the sorting of viral and endogenous glycoproteins in epithelial cells and hippocampal neurons suggest that there may be similarities in the mechanism of sorting in these two cell types^{1–3}. The mechanisms that maintain the distinct composition of the two plasma membrane domains in these two cell types must, however, be different. We have proposed the existence of a functional barrier at the axonal hillock/initial segment which prevents the intermixing of membrane constituents^{1,2}. Here we test this hypothesis by fusing liposomes containing fluorescent phospholipids into the plasma membrane of polarized hippocampal cells in culture. Fusion was induced by lowering the pH and mediated by influenza virus haemagglutinin expressed on the axonal surface of neurons infected with fowl plague virus. Labelling was found exclusively on axons after fusion. Although the fused lipids were mobile on the axonal membrane, no labelling was detected on the cell body and dendritic surfaces. These results suggest that there is a diffusion barrier at the axonal hillock/initial segment which maintains the compositional differences between the axonal and somatodendritic domains.

We used the same experimental strategy as that used to study

the fence function of the tight junctions in MDCK cells^{4,5}. Thus fluorescently labelled water-insoluble lipids were inserted into the plasma membrane of fowl plague virus (FPV)-infected hippocampal cells by exploiting the fusogenic activity of the influenza virus spike glycoprotein haemagglutinin (HA)^{4–7}. Liposomes containing the fluorescent lipids *N*-(lissamine rhodamine B sulphonyl)-dioleoylphosphatidylethanolamine (*N*-Rh-DOPE), *N*-(7-nitro 2,1,3-benzoxadiazol-4-yl)dioleoylphosphatidylethanolamine (*N*-NBD-DOPE), and the glycolipid IV³NeuAcII³NeuAcα-GgOseCer (G_{D1a}) were added to hippocampal neurons in culture immediately after FPV infection and the fluorescence pattern analysed by light microscopy (Fig. 1). After incubation with the labelled liposomes for 30 min at 7 °C, only rhodamine labelling was seen. Because of resonance energy transfer of NBD to rhodamine^{8,9}, little or no NBD signal was detected, suggesting that under these conditions adhesion of the liposomes to the plasma membrane but not fusion was the predominant event. The rhodamine labelling was exclusively present along the axons of infected cells (Fig. 1a–c). No rhodamine labelling was seen on the surface of cell bodies or dendrites.

Fusion was next induced by lowering the pH of the incubation medium. Immediate microscopic analysis revealed intense fluorescence along the axon with both rhodamine and NBD (Fig. 1d–f). The fact that NBD fluorescence could be visualized suggests that efficient fusion of the liposomes with the plasma membrane had taken place (as the separation of the fluorochromes takes place after fusion, energy transfer can no longer occur^{8,9}). The labelling with rhodamine and NBD had an identical distribution: exclusively on the axon and abruptly interrupted at the point where the axon emerged from the cell body or the base of a dendrite close to the cell body. As only 20% of the axons were labelled, the cell body or dendrite from which each labelled axon had originated could be unequivocally traced. The labelled axonal surface showed characteristic bulbous enlargements connected by segments of normal diameter. These bulbs are probably due to an increase in the plasma membrane surface area after fusion of liposomes. The labelling

* Present address: Department of Physics, Faculty of Science, Tohoku University, Aoba, Aramaki, Aobaku, Sendai 980, Japan.

‡ To whom correspondence should be addressed.

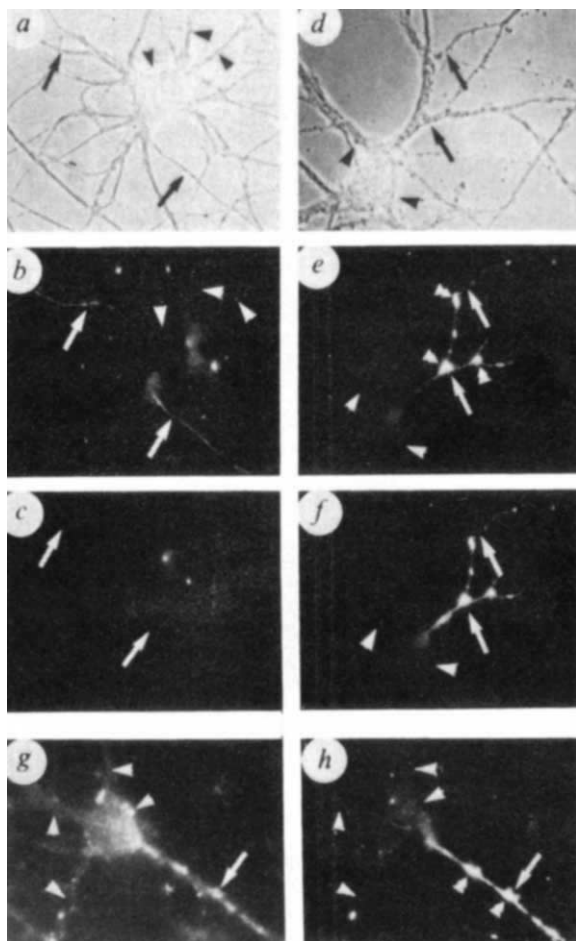


FIG. 1 Adhesion (a–c) and fusion (d–f) of *N*-NBD-DOPE- and *N*-Rh-DOPE-labelled liposomes to FPV-infected hippocampal cells. After adhesion prominent rhodamine fluorescence was observed (b) whereas NBD emission (c) is almost invisible, indicating that resonance energy has been transferred from NBD to rhodamine^{8,9}. The only labelled structures are axons (arrowed) as judged from their typical morphological appearance (thin and of uniform diameter). The cell body and dendrites, visible in the corresponding phase contrast photograph (a) appear to be unlabelled (arrowheads). After fusion, exclusive axonal labelling (arrowed) is seen for NBD (f) and rhodamine (e) (no energy transfer has occurred). The labelling begins at the axonal origin from the cell body. Axonal labelling is not only restricted to the first 30–50 μ m shown here but is also present along the entire axonal surface (not shown). By comparison with the phase contrast image (d) the absence of labelling of the cell body and dendrites is evident (arrowheads). g–h, fusion of liposomes containing *N*-Rh-DOPE and the water-soluble molecule calcein. Rhodamine labelling is only detectable along the axon (arrow in h), whereas calcein also labels the cell body and the dendrites (arrowheads in g). Note in e and h the bulbous enlargements of the axon (double arrowheads). Scale bar, 10 μ m.

METHODS. Small unilamellar liposomes were prepared by ethanol injection²³ from 93 mol% of dioleoylphosphatidylcholine (DOPC), one mol% each of the non-exchangeable fluorescent lipids *N*-Rh-DOPE and *N*-NBD-DOPE as described⁸ and 5 mol% of the glycolipid G_{D1a} (ref. 4). Liposomes containing the water-soluble marker calcein (Sigma) were prepared according to ref. 18. DOPC, *N*-Rh-DOPE and *N*-NBD-DOPE were purchased from Avanti Polar Lipids (Birmingham, AL); G_{D1a} was from Supelco (Bellefonte, PA). Hippocampal cells were cultured on glass coverslips from 18-day-old rat embryos as described²⁴. After 2 weeks in culture, cells were infected for 9–12 h with FPV as before¹. Binding and fusion of the labelled liposomes to the plasma membrane of the infected neurons were done according to ref. 4. After infection the coverslips were rinsed briefly in Hank's balanced salt solution (HBSS) and incubated for 30 min at 7 °C with 1 mM labelled liposomes followed by a 1 min incubation at 37 °C at pH 5. At low temperature only adhesion occurs, whereas the high temperature and low pH induced liposome-membrane fusion. After liposome incubation, coverslips were rinsed several times in ice-cold HBSS and mounted on a depression slide and the specimen was observed immediately in the fluorescence microscope.

FIG. 2 Distribution of labelled lipids after fusing asymmetrically labelled liposomes to FPV infected neurons. a, Addition of the quenching agent dithionite ('D') to symmetrically labelled liposomes results in a 50% loss of fluorescence. Fluorescence is totally lost after the addition of the detergent Triton X-100 ('T'). c, Fusion of *N*-NBD-DOPE-labelled liposomes in which the outer leaflet lipids were quenched by dithionite before addition to the cells (as in a) reveals exclusive axonal labelling. The labelling is interrupted at the axonal origin, which in this cell occurs from a dendrite (double arrowheads). The cell body and dendrites, visible in the corresponding phase contrast photograph (b, arrowheads), are not labelled. Some fluorescence near the cell body may correspond to cell debris. d–e Addition of the quenching agent to the cells after fusing symmetrically labelled liposomes also reveals the exclusive localization of fluorescence to the axon (arrows in e) without any evident signal in the cell body or dendrites (arrowheads in d). Note in c and e the bulbous enlargements of the axonal diameter. The efficiency of the quenching procedure in living cells is shown in f and g. After binding of *N*-NBD-DOPE to the neuronal surface, the entire cell architecture becomes fluorescent (f). After addition of dithionite all labelling disappears (g). Scale bar, 10 μ m.

METHODS. Liposomes containing 1 mol% *N*-NBD-DOPE, 5 mol% G_{D1a} and 94 mol% DOPC were made as described in Fig. 1. Outer leaflet lipids were quenched by the addition of 30 mM sodium dithionite (Sigma) to the liposome solution during constant mixing²⁵. The time course of fluorescence emission was followed (a). Free dithionite was removed by gel filtration chromatography on Sephadex G-50 and these asymmetric liposomes were used for the experiments shown in b and c. After fusion to the cells of symmetrically labelled liposomes dithionite was added to the culture medium at a concentration of 88 mM (experiment shown in d and e). Control bleaching is efficient: *N*-NBD-DOPE in ethanol was added to the cells (final concentration 14 μ M) at low temperature (4 °C, 5 min) then quenched by adding 88 mM dithionite (f, g).

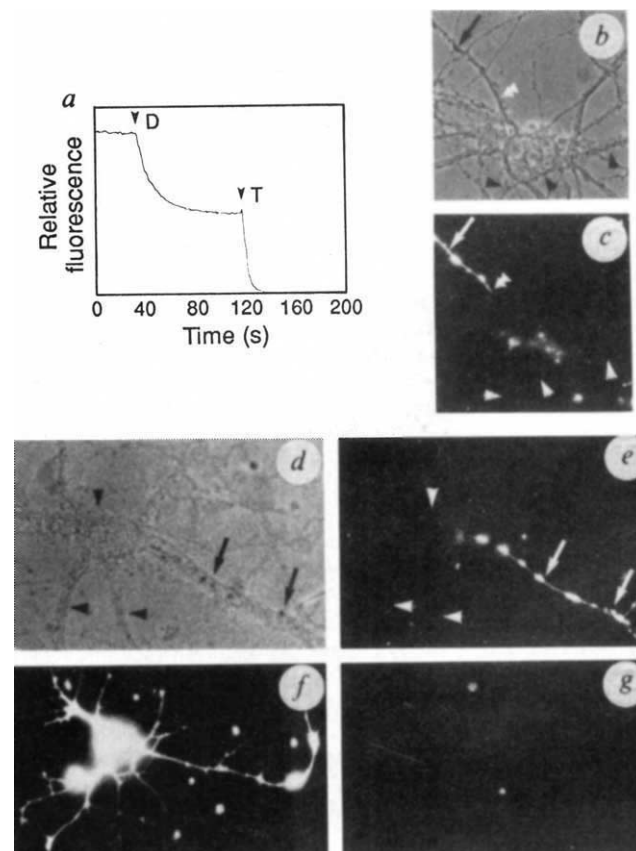
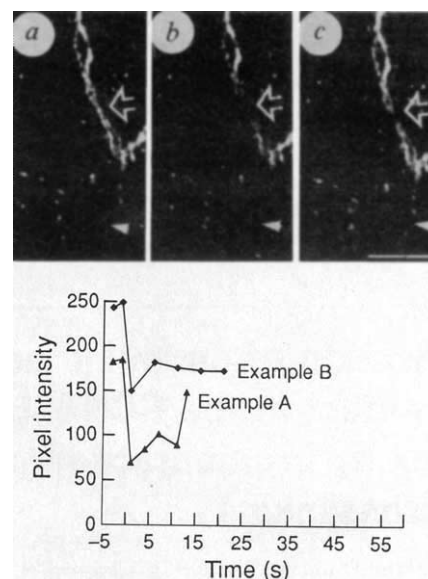


FIG. 3 Photobleaching of NBD lipid fluorescence is followed by rapid recovery in a mature neuron infected with FPV. The fluorescence photographs show the appearance of NBD-labelling in an axon before (a), and 1 s (b) and 5 s (c) after bleaching. The site of bleaching is indicated by an arrow. In this cell (which corresponds to example B in the accompanying graph) a 50% decrease of fluorescence was observed 1 s after bleaching. 40% of fluorescence recovered in the subsequent 5 s. Note in this confocal image the lack of labelling in the cell body and dendrites (arrowheads). The graph also shows example A in which fluorescence recovery was more dramatic (70% in 12 s). Photobleaching of axonal fluorescence after liposome-membrane adhesion did not result in fluorescence recovery (not shown). Scale bar, 10 μ m.

METHODS. Cultured hippocampal neurons were infected with FPV, fused with *N*-NBD-DOPE- and *N*-Rh-DOPE-labelled liposomes, and mounted on a depression slide as described in Fig. 1. The photobleaching and all microscopy were done with a Compact Confocal Microscope (EMBL, Heidelberg, Germany) fitted with a Zeiss 100 $\times$, 1.4 NA infinity corrected plan apochromat objective. Infected neurons were first located using conventional optics, a mercury light source, and rhodamine filters. Subsequent microscopy was in confocal mode with an argon laser light source. Axonal focal plane was established under rhodamine excitation conditions. NBD fluorescence was excited with the 488 nm laser line, 488a filter combination, power level 0.5 W. For photobleaching the axon was bisected for 300 ms excitation period spread along a 5 μ m line. Images were collected in a 60 $\times$ 60 μ m field size with each image being the average of two passes of the laser across the field. Images were collected under automated software control. (Further details of bleach procedure and instrument control will be published elsewhere (B.S., Hanninen, P., Stelzer, E., and Kreis, T.)) The digital images resulting from the confocal microscopy were smoothed and noise reduced with NIH Image 1.43 on a Macintosh Ci computer. For quantitative analysis the average pixel intensity in a 5 $\times$ 5 block in the centre of the bleached area was determined. No change in fluorescence intensity was seen in a control, non-bleached region of the cell during the image collection process.



in the axonal initial segment did not disappear after prolonged times of observation (30–60 min); the cell body and dendritic plasma membrane always appeared unlabelled even at these longer times after fusion. The axonal labelling was haemagglutinin-liposome specific as neither uninfected nor infected cells incubated with liposomes lacking the HA-receptor G_{D1a} showed any detectable labelling (results not shown). In stage-3 cells, in which there is morphological polarity^{10,11} but the HA glycoprotein is delivered to both axons and dendrites¹, binding and fusion of the labelled liposomes occurred over the entire neuronal surface (results not shown). These results suggest that the labelled lipids are trapped in the axon of mature neurons and cannot diffuse into the cell body owing to the existence of a physical barrier between these two territories. To show that the barrier was not located in the axonal cytoplasmic lumen, we fused liposomes containing the fluorescent water-soluble marker calcein. After fusion the entire cell cytosol in the axon, cell body and dendrites was labelled with calcein (Fig. 1g–h).

We next analysed whether the functional barrier was located in the inner and outer leaflets of the axonal membrane by using symmetrically and asymmetrically labelled liposomes (Fig. 2). Addition of liposomes to the infected neurons containing labelled lipids either in both leaflets or exclusively in the inner leaflet, as demonstrated by spectrofluorometric analysis (Fig. 2a), showed the typical continuous axonal labelling abruptly interrupted at the axonal origin (Fig. 2b, c). Similarly, no cell body or dendritic labelling was seen when symmetrically labelled liposomes were fused to the cells and outer leaflet lipids quenched after fusion (Fig. 2d–e). This result suggests that in neurons the diffusion barrier is located in both plasma membrane leaflets, unlike in MDCK cells where lipids in the inner leaflet can diffuse freely⁵.

To confirm that the fluorescent lipids were dispersed in the axon membrane, and to demonstrate by a second means the existence of the fusion event at the initial segment, we used

photobleaching followed by fluorescence recovery (Fig. 3). FPV-infected neurons were incubated with *N*-NBD and *N*-Rh co-labelled liposomes and an area $\sim 0.6 \mu$ m wide across the axon, close to its emergence from the cell body, was bleached using a laser beam. We found 40% recovery of fluorescence 5 s after bleaching in four independent determinations.

We know very little about how the differences in molecular organization of axons and dendrites in polarized neurons are generated and maintained. In polarized epithelia, restriction of membrane proteins and lipids to the apical and basolateral domains is accomplished by the tight junctions^{5,12} and by interactions with the membrane cytoskeletal proteins fodrin and ankyrin^{13,14}. Although tight junctions in neurons are not present, our results show that a fence to lipid movement is present at the junction between the axon and cell perikaryon. Fluorescently labelled liposomes only bind and fuse to the axon of FPV-infected neurons. The fused lipids were mobile in the axonal membrane and do not diffuse to the cell body and dendritic surfaces. Furthermore, the intensity of the axonal labelling did not decrease after prolonged observation. We have also demonstrated that the functional barrier extends across the cytoplasmic and exoplasmic leaflets of the membrane. It is possible that in neurons, as in other non-neuronal cells, ultrastructural specializations of the plasma membrane may determine the existence of a tight junction-like barrier. For example, *Chlamydomonas reinhardtii* has a barrier that prevents the diffusion of cell body agglutinins to the flagella¹⁵. Mammalian sperm cells also have a diffusion barrier which restricts the glycerophosphatidylinositol-anchored protein PH-20 to the posterior head of the acrosome¹⁶. In both cases characteristic membrane specializations are present at the sites of diffusion barriers^{17–19}. In neurons, the plasma membrane at the axonal hillock/initial segment shows specializations, such as the layer of dense granular material beneath the plasma membrane and the clustering of Na^+ channels^{20–22}, which may be the structural elements underlying the functional barrier shown here. □

Received 1 June; accepted 4 September 1992.

1. Dotti, C. G. & Simons, K. *Cell* **62**, 63–72 (1990).
2. Dotti, C. G., Parton, R. G. & Simons, K. *Nature* **349**, 158–161 (1991).
3. Powell, S., Cunningham, B., Edelman, G. & Rodriguez Boulan, E. *Nature* **353**, 76–77 (1991).
4. van Meer, G., Davoust, J. & Simons, K. *Biochemistry* **24**, 3593–3602 (1985).
5. van Meer, G. & Simons, K. *EMBO J.* **5**, 1455–1464 (1986).
6. White, J., Matlin, K. & Helenius, A. *J. Cell Biol.* **89**, 674–679 (1981).
7. White, J., Helenius, A. & Kartenbeck, J. *EMBO J.* **1**, 217–222 (1982).
8. Kobayashi, T. & Pagano, R. *Cell* **55**, 797–805 (1988).
9. Pagano, R. E. *Meth. Cell Biol.* **29**, 75–81 (1989).
10. Dotti, C., Sullivan, C. & Banker, G. *J. Neurosci.* **8**, 1454–1468 (1988).
11. Dotti, C. & Banker, G. *Nature* **330**, 254–256 (1987).
12. Gumbiner, B. *Am. J. Physiol.* **253**, C749–C758 (1987).
13. Nelson, W. J. & Hammerton, R. W. *J. Cell Biol.* **108**, 893–902 (1989).
14. Nelson, W. J. & Veshnock, C. K. *Nature* **328**, 533–536 (1987).
15. Hunnicutt, G., Kosfisz, M. & Snell, W. *J. Cell Biol.* **111**, 1605–1616 (1990).

16. Cowan, A., Myles, D. & Koppel, D. *J. Cell Biol.* **104**, 917–923 (1987).
17. Gilula, N. & Satir, P. *J. Cell Biol.* **53**, 494–509 (1972).
18. Weiss, R., Goodenough, D. & Goodenough, U. *J. Cell Biol.* **72**, 133–143 (1977).
19. Friend, D. & Fawcett, D. *J. Cell Biol.* **63**, 641–664 (1974).
20. Peters, A., Palay, S. & Webster, H. *The Fine Structure of the Nervous System* (Oxford Univ. Press, New York, 1991).
21. Angelides, K., Elmer, L., Loftus, D. & Elson, E. *J. Cell Biol.* **106**, 1911–1925 (1988).
22. Srinivasan, Y., Elmer, L., Davis, J., Bennett, V. & Angelides, K. *Nature* **333**, 177–180 (1988).
23. Kremer, J., v.d. Esker, M., Pathmamanoharan, C. & Wiersema, P. *Biochemistry* **16**, 3932–3935 (1977).
24. Banker, G. & Goslin, K. *Culturing Nerve Cells* (MIT Press, Cambridge, 1991).
25. McIntyre, J. C. & Sleight, R. G. *Biochemistry* **30**, 11819–11827 (1991).

ACKNOWLEDGEMENTS. We thank L. Meyn for preparing the hippocampal culture and R. Parton for comments on the manuscript. B.S. holds a Fogarty Senior International Fellowship from the NIH. T.K. was supported by a Human Frontier Science Program long-term fellowship.

Identification of an amino acid–base contact in the GCN4–DNA complex by bromouracil-mediated photocrosslinking

Erich E. Blatter, Yon W. Ebright & Richard H. Ebright*

Department of Chemistry and Waksman Institute, Rutgers University, New Brunswick, New Jersey 08855, USA

*To whom correspondence should be addressed

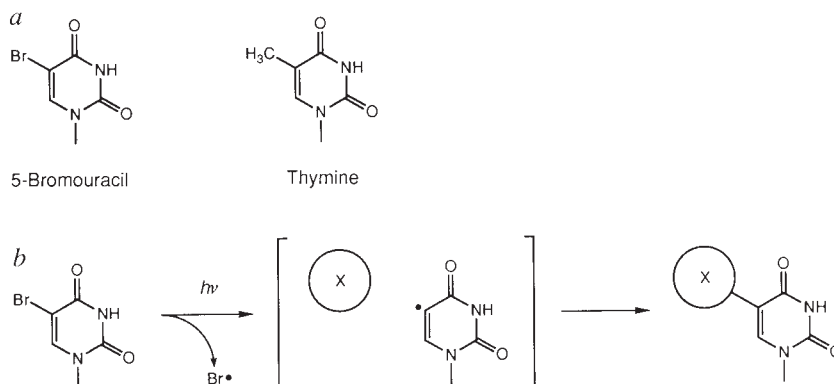
THE bZIP DNA-binding proteins are characterized by a 50-amino-acid DNA binding and dimerization motif, consisting of a highly basic DNA-binding region ('b') followed by a leucine zipper dimerization region ('ZIP')¹. The best characterized bZIP DNA-binding protein is GCN4, a yeast transcriptional activator^{2–6}. GCN4 binds to a 9-base-pair two-fold-symmetric DNA site, 5'-A₄T₃G₂A₁C₀T₁C₂A₃T₄-3' (refs 7–10). A detailed model known as the 'induced helical fork' model has been proposed for the structure of the GCN4–DNA complex⁴. Using a site-specific bromouracil-mediated photocrosslinking method, we show here that the alanine at position 238 of GCN4 contacts, or is close to, the thymine 5-methyl of A·T at position +3 of the DNA site in the GCN4–DNA complex. Our results strongly support the induced helical fork model⁴. Our site-specific bromouracil-mediated photocrosslinking method requires no prior information regarding the structure of the protein or the structure of the protein–DNA complex and should be generalizable to DNA-binding proteins that interact with the DNA major groove^{11,12}.

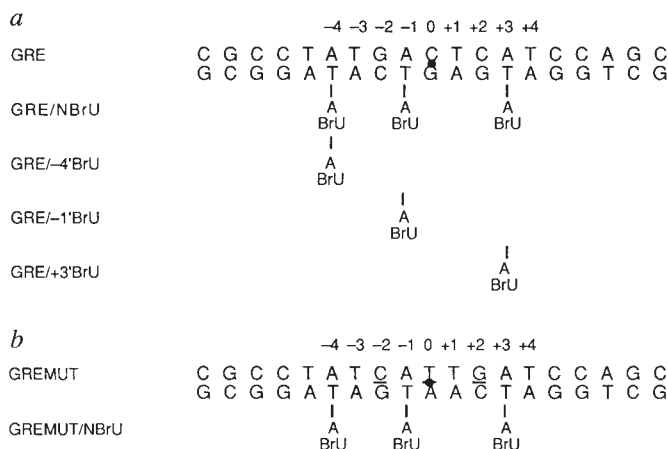
5-Bromouracil (BrU) is an isosteric photoreactive analogue of thymine (Fig. 1)¹³. We have incorporated BrU at a single position within the DNA site for GCN4, prepared the protein–DNA complex, irradiated the protein–DNA complex with ultraviolet light and determined the amino acid at which crosslinking

occurs. We used a 58-amino-acid protein fragment, GCN4p, corresponding to the bZIP motif of GCN4 (amino acids 226 to 281 of GCN4)^{14,15}; GCN4p contains all determinants of GCN4 for DNA binding and for dimerization¹⁴. We also used a set of seven 19-base-pair DNA fragments (Fig. 2). DNA fragments GRE, GRE/NBrU, GRE/–4'BrU, GRE/–1'BrU and GRE/+3'BrU contain the DNA site for GCN4 (Fig. 2a). In DNA fragment GRE/NBrU, every thymine of the bottom strand of the DNA site is replaced by BrU. In DNA fragments GRE/–4'BrU, GRE/–1'BrU, and GRE/+3'BrU, a single thymine of the bottom strand is replaced by BrU. Gel-mobility-shift DNA-binding experiments^{16,17} established that DNA fragments GRE, GRE/NBrU, GRE/–4'BrU, GRE/–1'BrU, and GRE/+3'BrU are fully functional in specific GCN4p–DNA complex formation ($K_{obs} \approx 7 \times 10^8 \text{ M}^{-1}$). DNA fragments GREUT and GREUT/NBrU contain a 'mutant' derivative of the DNA site for GCN4 (Fig. 2b). In DNA fragment GREUT/NBrU, every thymine of the bottom strand of the mutant derivative of the DNA site for GCN4 is replaced by BrU. Gel-mobility-shift DNA-binding experiments^{16,17} established that DNA fragments GREUT and GREUT/NBrU are not functional in specific GCN4p–DNA complex formation.

To determine whether DNA fragment GRE/NBrU can crosslink with GCN4p in the GCN4p–DNA complex, we formed the GCN4p–DNA complex, irradiated it with ultraviolet light and analysed the reaction products by polyacrylamide gel electrophoresis under denaturing conditions. The results (Fig. 3a) indicate that ultraviolet irradiation yields a crosslinked GCN4p–DNA complex, the efficiency of crosslinking being ~2%. Control experiments established that ultraviolet irradiation, GCN4p and BrU-containing DNA are necessary for crosslinking. Additional control experiments established that specific GCN4p–DNA complex formation is necessary for crosslinking. Thus, crosslinking is inhibited by excess DNA fragment GRE but not by excess DNA fragment GREUT. Furthermore, crosslinking does not occur with DNA fragment GREUT/NBrU. We conclude that at least one BrU residue in DNA fragment

FIG. 1 a, Structures of 5-bromouracil ('BrU') and thymine. BrU is an isosteric photoreactive analogue of thymine (see ref. 13 for a review). Most sequence-specific DNA-binding proteins cannot distinguish between BrU and thymine^{22–26}. BrU can be introduced at any position(s) within a synthetic oligodeoxynucleotide using 5-bromo-2'-deoxyuridine- β -cyanoethylphosphoramidite^{25,26}. b, Photochemistry of BrU (refs 11, 12, 27–29). In the dark BrU is unreactive. But on irradiation with ultraviolet light, bond homolysis occurs and a highly reactive uracyl radical is generated. If there is organic material in direct van der Waals contact with the uracyl radical, covalent bond formation may occur.





GRE/NBrU can crosslink with GCN4p and that the 5-methyl of at least one thymine of the bottom strand of the DNA site contacts, or is close to, GCN4p.

To determine which BrU residue(s) of DNA fragment GRE/NBrU can crosslink with GCN4p, we did analogous experiments using DNA fragments GRE/-4'BrU, GRE/-1'BrU and GRE/+3'BrU, each of which has a single thymine of the bottom strand of the DNA site replaced by BrU. The results shown in Fig. 3b indicate that there is little or no crosslinking with DNA fragments GRE/-4'BrU and GRE/-1'BrU, but that there is significant crosslinking with DNA fragment GRE/+3'BrU. The efficiency of crosslinking with DNA fragment GRE/+3'BrU is comparable to the efficiency of crosslinking with DNA fragment GRE/NBrU. We conclude that the bottom-strand BrU residue at position +3 of the DNA site is responsible for nearly all crosslinking with DNA fragment

FIG. 2 DNA fragments used in the analysis. (Positions within the DNA site for GCN4 are numbered as in refs 2 and 7-9; a different convention is used in ref. 4.) **a**, DNA fragments containing the DNA site for GCN4. In DNA fragment GRE/NBrU, every thymine of the bottom strand of the DNA site is replaced by BrU. In DNA fragments GRE/-4'BrU, GRE/-1'BrU and GRE/+3'BrU, a single thymine of the bottom strand of the DNA site is replaced by BrU. **b**, DNA fragments containing a 'mutant' derivative of the DNA site for GCN4. In DNA fragment GREMUT/NBrU, every thymine of the bottom strand of the mutant derivative of the DNA site for GCN4 is replaced by BrU.

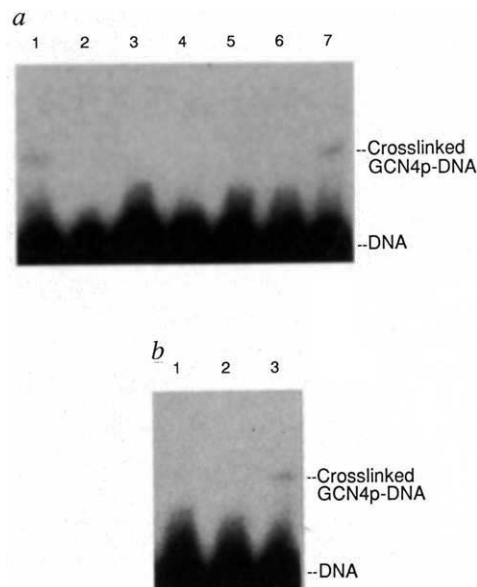
METHODS. For each DNA fragment, oligodeoxyribonucleotides corresponding to the top and bottom strands were synthesized using solid-phase β -cyanoethylphosphoramidite chemistry on an AB380A automated synthesizer. BrU was introduced using 5-bromo-2'-deoxyuridine- β -cyanoethylphosphoramidite (Glen Research, Inc.). Products were purified by trityl-on C18 reversed-phase HPLC, followed by detritylation and precipitation with ethanol. The oligodeoxyribonucleotide corresponding to the bottom strand was 5'-end-labelled as described³⁰ and the oligodeoxyribonucleotides corresponding to the top and bottom strands were annealed³⁰. All operations with BrU-containing oligodeoxyribonucleotides and DNA fragments were executed in darkness or under safe-lamp illumination.

GRE/NBrU, and that the 5-methyl of the bottom-strand thymine at position +3 of the DNA site contacts, or is close to, GCN4p.

To identify the amino acid(s) of GCN4p at which crosslinking occurs with DNA fragment GRE/+3'BrU, we prepared the crosslinked GCN4p-DNA complex in picomole quantities, digested it with trypsin, isolated the crosslinked peptide-DNA complex, and determined its amino-acid sequence. The sequence of the crosslinked peptide, Ala-Arg-Asn-Thr-Glu-X-Ala-Arg (Fig. 4A), corresponds to amino acids 233 to 240 of GCN4 (Fig. 4B), which are amino acids of the basic region of the bZIP motif of GCN4. No standard amino acid was identified at the sixth position of the crosslinked peptide. This position corresponds to Ala 238 of GCN4; therefore, we conclude that Ala 238 is where crosslinking occurs. The fact that amino-acid sequencing could proceed through amino-acid 238 indicates that crosslinking at amino-acid 238 does not involve the backbone

FIG. 3 Formation of crosslinked GCN4p-DNA complex. **a**, Data for DNA fragment GRE/NBrU. Lane 1, crosslinking reaction; lane 2, control reaction omitting ultraviolet irradiation; lane 3, control reaction omitting GCN4p; lane 4, control reaction using DNA fragment GRE in place of DNA fragment GRE/NBrU; lane 5, control reaction using DNA fragment GREMUT/NBrU in place of DNA fragment GRE/NBrU; lane 6, control reaction in the presence of a fivefold excess of DNA fragment GRE; lane 7, control reaction in the presence of a fivefold excess of DNA fragment GREMUT. **b**, Data for DNA fragments GRE/-4'BrU, GRE/-1'BrU and GRE/+3'BrU. Lane 1, crosslinking reaction using DNA fragment GRE/-4'BrU; lane 2, crosslinking reaction using DNA fragment GRE/-1'BrU; lane 3, crosslinking reaction using DNA fragment GRE/+3'BrU.

METHODS. GCN4p consists of two non-native amino acids, Met₁-Lys₂, followed by amino acids 226 to 281 of GCN4 (refs 14, 15). Plasmid pETPCR-1 encodes GCN4p under control of the bacteriophage T7 gene 10 promoter¹⁴ (gift from T. Ellenberger). GCN4p was expressed in *Escherichia coli* to a level of 50% total cell protein, and purified by precipitation with polyethylenimine and ammonium sulphate, followed by chromatography on phenyl-Sepharose and heparin-Sepharose¹⁴. GCN4p was $\geq 95\%$ homogeneous as assessed by SDS-polyacrylamide gel electrophoresis. Crosslinking reaction mixtures contained (in 50 μ l): 1.25 μ M GCN4p, 1.00 μ M ³²P-5'-end-labelled DNA fragment (200 Bq pmol⁻¹), 10 mM MOPS-NaOH (pH 7.3), 500 mM NaCl and 0.1 mM dithiothreitol. Reaction mixtures were incubated in the dark for 60 min at 22 °C, then ultraviolet-irradiated for 90 s at 22 °C (254 nm; 1,280 ergs min⁻² s⁻¹) in a Rayonet RPR100 photochemical reactor (Southern New England Ultraviolet). Under these conditions irradiation does not affect the specific activity of GCN4p. Irradiated reaction mixtures were precipitated with ethanol, redissolved in 10 μ l 5 M urea, 50 mM NaOH, 0.5 mM EDTA, 0.025% bromophenol blue and 0.025% xylene cyanol, and electrophoresed on 10% polyacrylamide-8 M urea slab gels. After electrophoresis, gels were dried and autoradiographed on to Kodak X-OMAT XARS film. Efficiency of crosslinking was determined by Cerenkov counting of excised bands.



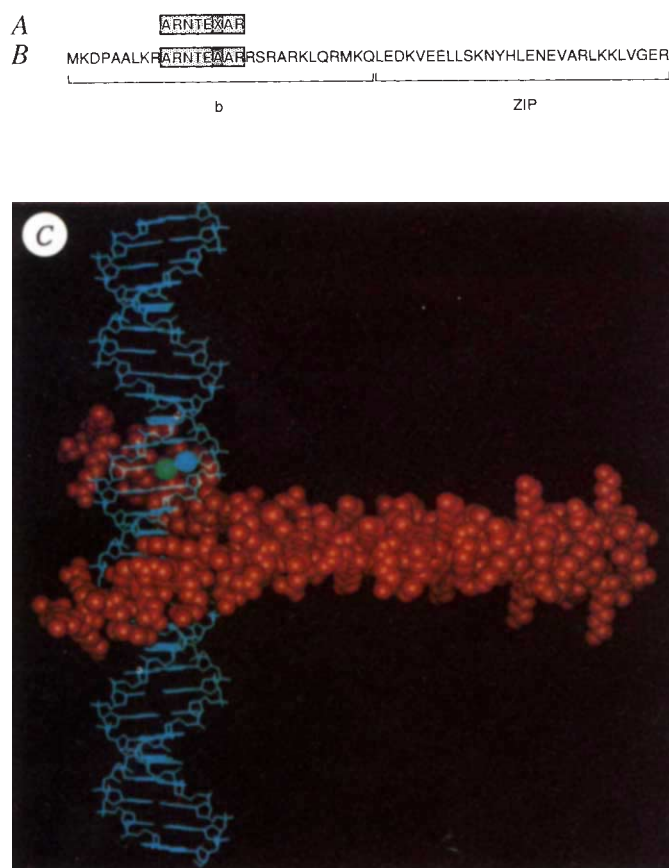


FIG. 4 **A**, Amino-acid sequence of the crosslinked peptide. At the sixth position of the crosslinked peptide, no standard amino acid was identified by sequencing ('X'). **B**, Amino-acid sequence of GCN4p (refs 14, 15). The basic region of GCN4p ('b') and the leucine zipper region of GCN4p ('ZIP') are indicated. **C**, Induced helical fork model for the structure of the GCN4-DNA complex⁴. The van der Waals surface of the C β atom of Ala 238 of one subunit of GCN4 is in green, and the van der Waals surface of the 5-methyl of the bottom-strand thymine at position +3 of the DNA site is in blue. Coordinates for the model were obtained from W. DeGrado and R. Hoess. **METHODS**. Crosslinked GCN4p-DNA complex was prepared using DNA fragment GRE/+3BrU as described for Fig. 3, except that the reaction volume was 100 ml (ultraviolet-irradiated in 25-ml aliquots) and the specific activity of the DNA fragment was 10 Bq pmol⁻¹. Crosslinked GCN4p-DNA complex was purified from uncrosslinked DNA by extraction into phenol, then recovered from phenol by back-extraction into water, and proteolytically digested with trypsin. The irradiated sample was adjusted to 0.1% SDS, incubated for 5 min at 70 °C and extracted with 25 ml Tris-saturated phenol. The resulting phenol phase was back-extracted with 6 ml water plus 25 ml ether. The resulting aqueous phase was washed twice with 25 ml ether and lyophilized. The sample was redissolved in 400 μ l 400 mM ammonium bicarbonate (pH 8.8) and 8 M urea. After incubation for 1 min at 70 °C, 1.2 ml water and 0.5 μ g sequencing-grade modified trypsin (Promega) was added and the sample incubated for 2 h at 37 °C. The resulting crosslinked peptide-DNA complex was purified from uncrosslinked peptides by denaturing polyacrylamide gel electrophoresis. The sample was ethanol-precipitated twice, redissolved in 20 μ l 6 M urea, 0.025% bromophenol blue and 0.024% xylene cyanol, and electrophoresed through a 10% polyacrylamide-2 M urea slab gel (pre-run for 30 min with cathode buffer containing 1 mM thioglycolic acid). The crosslinked peptide-DNA complex was excised from the gel and eluted into 1 ml 10 mM ammonium bicarbonate (pH 8.8) by shaking for 15 h at 37 °C. Crosslinked peptide-DNA complex was lyophilized, redissolved in 100 μ l 10 mM ammonium bicarbonate (pH 8.8), applied to a 5-ml column of Bio-Gel P-6DG (Bio-Rad), eluted in the same buffer and lyophilized; recovery was 80 pmol. The amino-acid sequence of the crosslinked peptide was determined by D. Coiffe using an AB477A automated gas-phase protein sequencer.

amide nitrogen atom, the backbone carbonyl carbon atom or the backbone carbonyl oxygen atom. We conclude that the backbone C α atom or the side-chain C β atom of Ala 238 of GCN4 is the site of crosslinking.

Our results establish that the C α or C β atom of Ala 238 of GCN4 contacts, or is close to, the 5-methyl of the bottom-strand thymine at position +3 of the DNA site in solution. This is in striking agreement with the 'induced helical fork' model for the structure of the GCN4-DNA complex⁴, which predicts that the C β atom of Ala 238 of GCN4 is in van der Waals contact with the 5-methyl group of the bottom-strand thymine at position +3 of the DNA site in solution (Fig. 4C). This is also in striking agreement with the crystallographic structure of the GCN4-DNA complex (T. Ellenberger and S. Harrison, personal communication).

Contacts with the thymine 5-methyl can be an important, determinant of specificity for thymine in protein-DNA interac-

tion¹⁸⁻²⁰. We propose that a van der Waals contact between the C α or C β atom of Ala 238 of GCN4 and the 5-methyl of the bottom-strand thymine at position +3 of the DNA site is part, or all, of the basis of specificity for A-T at position +3 of the DNA site. This proposal is in agreement with the observations that removal of the 5-methyl of the bottom-strand thymine at position +3 of the DNA site decreases affinity in GCN4-DNA complex formation²¹ and that substitution of Ala 238 of GCN4 affects the affinity and specificity of complex formation (K. Struhl, personal communication).

Fos and Jun are two other bZIP DNA-binding proteins^{1,2,4}. Fos and Jun share sequence homology with GCN4, including Ala 238 of GCN4, and recognize the same DNA site as GCN4 (refs 2, 4). We are currently using site-specific BrU-mediated photocrosslinking to determine whether equivalent amino acids in GCN4, Fos and Jun make similar contacts in their respective complexes with DNA. □

Received 21 May; accepted 3 September 1992.

- Vinson, C., Sigler, P. & McKnight, S. *Science* **246**, 911-916 (1989).
- Struhl, K. in *Transcriptional Regulation* (eds Yamamoto, K. & McKnight, S.) (Cold Spring Harbor Laboratory Press, New York, in the press).
- Talanian, R., McKnight, C. J. & Kim, P. *Science* **249**, 769-771 (1990).
- O'Neill, K., Hoess, R. & DeGrado, W. *Science* **249**, 774-778 (1990).
- O'Shea, E., Klemm, J., Kim, P. & Alber, T. *Science* **254**, 539-544 (1991).
- Saudek, V. et al. *Biochemistry* **30**, 1310-1317 (1991).
- Hill, D., Hope, I., Macke, J. & Struhl, K. *Science* **234**, 451-457 (1986).
- Oliphant, A., Brandl, C. & Struhl, K. *Molec. cell. Biol.* **9**, 2944-2949 (1989).
- Mavrothalassitis, G., Beal, G. & Papas, T. *DNA Cell Biol.* **9**, 783-788 (1990).
- Gartenberg, M., Ampe, C., Steitz, T. & Crothers, D. *Proc. natn. Acad. Sci. U.S.A.* **87**, 6034-6038 (1990).
- Wick, K. & Matthews, K. *J. biol. Chem.* **266**, 6106-6112 (1991).
- Allen, T., Wick, K. & Matthews, K. *J. biol. Chem.* **266**, 6113-6119 (1991).
- Hutchinson, F. Q. *Rev. Biophys.* **6**, 201-246 (1973).
- Weiss, M. et al. *Nature* **347**, 575-578 (1990).
- Weiss, M. *Biochemistry* **29**, 8020-8024 (1990).
- Fried, M. & Crothers, D. *Nucleic Acids Res.* **9**, 6505-6525 (1981).
- Garner, M. & Revzin, A. *Nucleic Acids Res.* **9**, 3047-3060 (1981).
- Fisher, E. & Caruthers, M. *Nucleic Acids Res.* **7**, 401-416 (1979).
- Wharton, R. & Ptashne, M. *Nature* **326**, 888-891 (1987).
- Gunasekera, A., Ebright, Y. & Ebright, R. *J. biol. Chem.* **267**, 14713-14720 (1992).
- Pu, W. & Struhl, K. *Nucleic Acids Res.* **20**, 771-775 (1992).
- Goeddel, D., Yansura, D., Winston, C. & Caruthers, M. *J. molec. Biol.* **123**, 661-687 (1978).
- Brennan, C., van Cleve, M. & Gumpert, R. *J. biol. Chem.* **261**, 7270-7278 (1986).
- Hayakawa, T., Ono, A. & Ueda, T. *Nucleic Acids Res.* **16**, 4761-4776 (1988).
- Fliess, A., Wolfes, H. & Pingoud, A. *Nucleic Acids Res.* **16**, 11781-11793 (1988).
- Risse, G., Jooss, K., Neuberg, M., Brüller, H.-J. & Müller, R. *EMBO J.* **8**, 3825-3832 (1989).
- Lin, S.-Y. & Riggs, A. *Proc. natn. Acad. Sci. U.S.A.* **71**, 947-951 (1974).
- Ogata, R. & Gilbert, W. *Proc. natn. Acad. Sci. U.S.A.* **74**, 4973-4976 (1977).
- Simpson, R. *Cell* **18**, 277-285 (1979).
- Ebright, R., Ebright, Y. & Gunasekera, A. *Nucleic Acids Res.* **17**, 10295-10305 (1989).

ACKNOWLEDGEMENTS. We thank T. Ellenberger for the plasmid encoding GCN4p and for discussion; W. DeGrado and R. Hoess for coordinates for the induced helical fork model; and B. Haley and R. Evans for discussion. This work was supported by a Searle Scholar Award to R.H.E. and by a Merck Postdoctoral Fellowship to E.E.B.

Short alanine-based peptides may form 3_{10} -helices and not α -helices in aqueous solution

Siobhan M. Miick, Gary V. Martinez, Wayne R. Fiori,
A. Paul Todd & Glenn L. Millhauser*

Department of Chemistry and Biochemistry, University of California,
Santa Cruz, California 95064, USA

* To whom correspondence should be addressed

SHORT alanine peptides, containing 16 or 17 residues, appear to form α -helices in aqueous solution¹⁻⁴. But the main spectroscopic analyses used on helical peptides (circular dichroism⁵ and nuclear magnetic resonance⁶⁻⁸) cannot distinguish between an α -helix (in which the i th residue is hydrogen-bonded to residue $i+4$; ref. 9) and the next most common peptide helix, the 3_{10} -helix¹⁰ ($i \rightarrow i+3$ hydrogen-bonding). To address this problem we have designed single and doubly spin-labelled analogues of alanine-based peptides in which the nitroxide spin label forms an unbranched side chain extending from the sulphur atom of a cysteine residue. Here we report the circular dichroism, Fourier-transform infrared and electron-spin resonance spectra of these peptides under helix-forming conditions. The infrared absorbance gives an amide I' band with a frequency that is substantially different from that observed for α -helices. The electron-spin resonance spectra of doubly labelled helices show that the ranking of distances between side chains, around a single turn (residues 4-8), is inconsistent with an α -helical structure. Our experiments suggest that the more likely peptide geometry is a 3_{10} -helix.

The peptides used are analogues of the alanine-based 3K(I) peptide (which we refer to as 3K-1), which contains three lysine residues in an alanine chain with cysteine as position 1 (ref. 1, and Table 1). The cysteines are introduced to provide an attachment point for the nitroxide methanethiosulphonate spin label (MTSSL)^{11,12}. The introduction of a single spin-labelled Cys shifts the helix-coil equilibrium by only a few per cent. The circular dichroism spectra of the peptides at 1 °C are shown in

TABLE 1 Spin-labelled peptide sequences

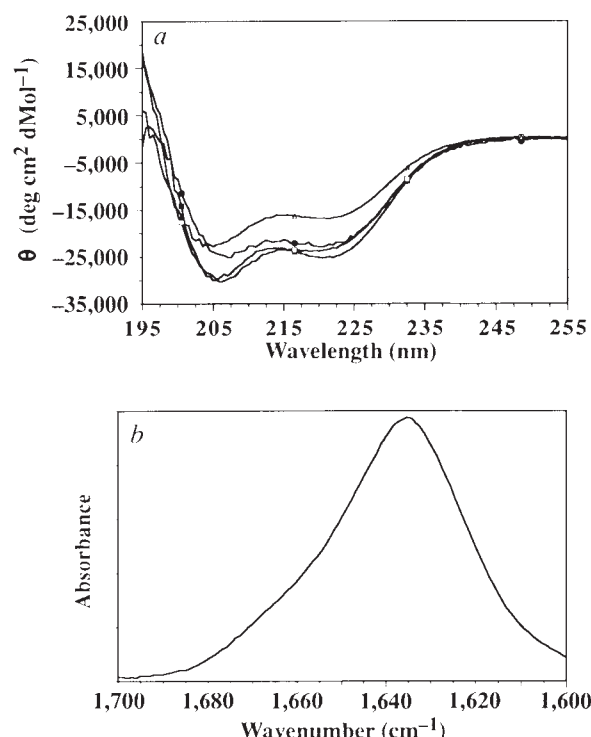
	Peptide sequence
Singly labelled	
3K-1	Ac-CAAAKAAAAKAAAAKA-NH ₂
3K-8	Ac-AAAAKAAKAAAAKA-NH ₂
Doubly labelled	
3K-(4, 6)	Ac-AAACKCAAAKAAAAKA-NH ₂
3K-(4, 7)	Ac-AAACKACA KAAAAKA-NH ₂
3K-(4, 8)	Ac-AAACKAACAKAAAAKA-NH ₂

Peptide sequences in single-letter amino-acid code. The 3K-8 peptide was custom-synthesized by the American Peptide Co. The remaining peptides were synthesized by Fmoc solid-phase peptide synthesis on a Rainin PS3 peptide synthesizer as carboxy-terminal amides using a 4-(2',4'-dimethoxyphenyl-Fmoc-aminomethyl)phenoxy resin and capped at the amino terminus with acetic anhydride. Cleavage from the resin was accomplished with a mixture of 90% trifluoroacetic acid (TFA), 5% thioanisole and 5% anisole, and the material was crystallized from ether. Initial purification was performed by gel filtration chromatography on a Sephadex G-10 column in 10 mM ammonium acetate, pH 4.7, and then by reverse-phase HPLC with a C18 resin using a gradient of 5-45% acetonitrile/H₂O in 0.1% TFA. The purified peptides were spin-labelled at the cysteine residue(s) by a disulphide linkage with MTSSL (Reanal, Budapest) in a 5% acetonitrile/0.1% TFA solution by addition of MTSSL in 10-fold excess. The reaction mixture was subsequently repurified by HPLC. The molecular masses of both the purified unlabelled and the purified labelled peptides were verified by fast-atom-bombardment mass spectrometry, and agreed with the expected value to within 1 AMU.

Fig. 1. As expected for Ala-based peptides, the circular dichroism signals of these spin-labelled analogues are strong. Both the 3K-(4, 7) and the 3K-(4, 8) are helical, whereas the 3K-(4, 6) shows a reduced circular dichroism signal.

Fourier-transform infrared (FTIR) spectra give information about the secondary structure of proteins¹³ and peptides¹⁴. For folded proteins in D₂O, α -helices are characterized by an amide I' band between 1,650 cm⁻¹ and 1,657 cm⁻¹. The FTIR spectrum for the 3K-1 analogue in D₂O at 1 °C is shown in Fig. 1. The band has a maximum at 1,637 cm⁻¹, which is uncharacteristic for an α -helix. As a comparison, polyaminoisobutyric acid forms 3_{10} helices in films¹⁵ and FTIR has shown that deuterated

FIG. 1 *a*, The circular dichroism spectra are shown with sample conditions of 1 °C in 5 mM MOPS buffer, pH 7.1, for the peptides 3K-8 (■), 3K-(4, 6) (△), 3K-(4, 7) (□), and 3K-(4, 8) (●) at peptide concentrations in the range of 50-150 μ M. The measurements were made on an Aviv 60DS spectropolarimeter in a 0.1-cm path-length cuvette. Peptide concentrations were determined with an accuracy of $\pm 3\%$ by double integration of ESR spectra and comparison to a 1 mM 4-hydroxy-TEMPO standard solution. *b*, The amide I' band of the FTIR spectrum of the 3K-1 peptide at 2 °C in 50 mM phosphate buffer and D₂O, pD 7, and ~ 1 mM peptide concentration. Peptide solutions were placed in liquid cell fitted with CaF₂ windows and a 50- μ m Teflon spacer. Spectra were obtained with a Nicolet 800 FTIR spectrometer continuously purged with N₂. The temperature was controlled using a cell jacket of circulating water. The spectrometer resolution was 2 cm⁻¹ and 1,024 interferograms were signal-averaged. The ratio of spectrum solvent with sample against the solvent spectrum was calculated and any residual water vapour peaks were subtracted from the spectrum. No additional smoothing or line narrowing was performed.



polyaminoisobutyric acid films give an amide I' band of $1,640\text{ cm}^{-1}$ (ref. 16), which is close to our result. Also, in a recent study of α -lactalbumin, an amide I' band at $1,639\text{ cm}^{-1}$ was assigned to a 3_{10} -helix¹⁷. The band we observe is broad and may indicate a superposition of helical conformations.

Electron-spin resonance (ESR) spectra of doubly labelled (biradical¹⁸) peptides allow us to rank the relative distances between side chains. As a qualitative guide for the expected interactions between side chains, Fig. 2a shows a helical wheel representation of an α -helix and a 3_{10} -helix with the relevant labelling positions. The distances between side-chain β -carbons $d_{\beta\beta}(i, j)$ for an α -helix and a 3_{10} -helix provide a quantitative ranking (Table 2) and show a distinct pattern for each of the two helical forms.

ESR spectra of the single labelled 3K-8 and of the three biradical peptides are shown in Fig. 2b. The spectra for the 3K-(4, 6) and the 3K-(4, 8) are very similar to the 3K-8, indicating a weak spin-spin interaction between the labels. But the 3K-(4, 7) shows substantial broadening. The peak-to-peak hyperfine line width δ of a biradical spectrum broadens as the distance between spins decreases^{19,20}. The reciprocals of the centre ($M_1 = 0$) spectral line widths δ^{-1} , which are proportional to the relative distances between nitroxides in the doubly labelled peptides, are reported in Table 2. Nitroxide-labelled side

chains are flexible so δ^{-1} represents an average distance between side chains. To investigate the consequence of multiple side-chain conformations, we have performed molecular dynamics calculations²¹ and found an average ranking of distances ($\langle d(i, j) \rangle$) between nitroxide N atoms (at positions i, j) that follows $\langle d(4, 7) \rangle \approx \langle d(4, 8) \rangle < \langle d(4, 6) \rangle$ for the α -helix and $\langle d(4, 7) \rangle < \langle d(4, 6) \rangle < \langle d(4, 8) \rangle$ for the 3_{10} -helix. This ranking matches that expected from $d_{\beta\beta}(i, j)$ values within experimental error. These ESR experiments and molecular dynamics calculations are summarized in Table 2. The δ^{-1} values for the 3K-(i, j) peptides show that $d(4, 7) < d(4, 6) < d(4, 8)$, matching the expected ranking for a 3_{10} -helix based on $d_{\beta\beta}(i, j)$ values and molecular dynamics calculations. The geometry of an α -helix ($d(4, 7) \approx d(4, 8) < d(4, 6)$) is clearly not observed.

The unfolded state of the peptide serves as a control experiment for double-label electron-spin resonance experiments. Guanidine hydrochloride fully denatures these Ala-based peptides¹, giving distance ranking of $d(4, 6) \approx d(4, 7) \approx d(4, 8)$, which matches neither of the helical configurations.

Trifluoroethanol at 20 mole per cent is believed to induce 100% helix formation in aqueous solution¹. The ESR spectra of the 3K-(4, 6) and 3K-(4, 8) analogues show very little change (data not shown). But the 3K-(4, 7) spectrum changes dramatically (Fig. 2b) and becomes a superposition of a typical three-

FIG. 2 a, Helical wheel representations for an α -helix and a 3_{10} -helix shown from an end-on view from the N terminus reveal the relative side-chain positions. The interaction patterns between the side chains of the doubly labelled peptides are distinct for the two conformations. The distances between β -carbons for the two geometries are given in Table 1. b, Continuous-wave ESR spectra of the 3K-8, 3K-(4, 6), 3K-(4, 7), and 3K-(4, 8) are shown for samples at 1°C and peptide concentrations of 1–1.8 mM in 5 mM MOPS, pH 7.1. Measurements were made on a Bruker ESP 380 equipped with a TE₁₀₂ rectangular cavity operating in the continuous wave mode with modulation amplitude 0.19 G, frequency 100 kHz and a 100 G scan-width. The spectra for these peptides do not vary in the concentration range 0.05–1.8 mM. The spectra for the 3K-(4, 7) in 20 mol per cent TFE were measured at 25°C at a concentration of 130 μM . Two interactions impact the ESR spectra of nitroxide biradicals (molecular species containing two localized unpaired electrons): overlap of two unpaired electron orbitals and dipole-dipole broadening between the two electron spins¹⁸. In the weak orbital overlap region, where the electrons are more distant from each other, these mechanisms mainly result in broadening of the three-line spectrum. But strong exchange spectra can appear as weak exchange if the spin-spin interaction is rapidly modulated. Careful spectral integration can reveal this²⁰. Integration of the absorption spectra of our doubly labelled peptides under denaturing conditions of increased temperature or exposure to guanidine hydrochloride shows that although the shape of the spectra can change dramatically, for example, giving five lines at high temperature, the signal integral never varies by more than a few per cent. Further, the narrow spectrum of the 3K-(4, 8) fits well with a lineshape composed of three Lorentzians. We conclude that all of our spectra are strictly weak exchange, except for the 3K-(4, 7) in TFE, which may be intermediate exchange.

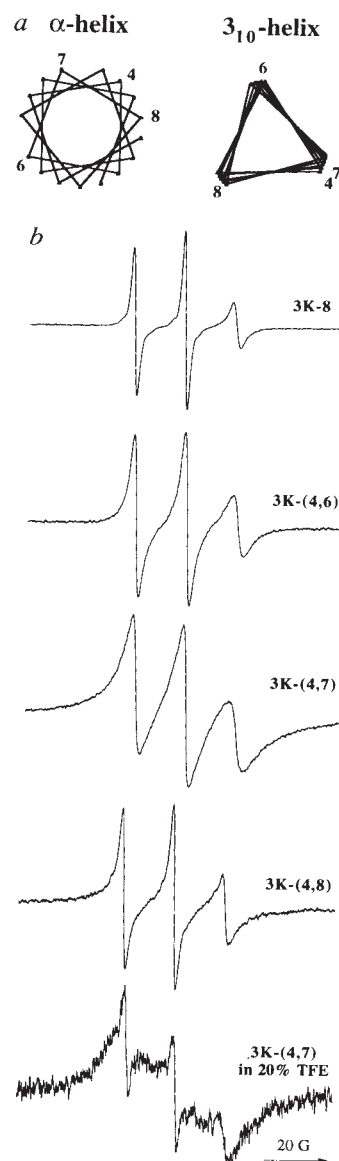


TABLE 2 Experimentally determined and calculated distances between labelled side chains

Distance between β -carbons for two helix geometries			
Helix	$d_{\beta\beta}(4, 6)$ (Å)	$d_{\beta\beta}(4, 7)$ (Å)	$d_{\beta\beta}(4, 8)$ (Å)
α -helix	7.68	6.56	6.79
3_{10} -helix	7.32	6.59	10.0
ESR results for the doubly labelled peptides			
Reciprocal linewidth ($M_i=0$)	3K-(4, 6)	3K-(4, 7)	3K-(4, 8)
$\delta^{-1}(\text{G}^{-1})$	0.410	0.353	0.621
Distance between nitroxide N atoms from molecular dynamics			
Helix	$\langle d(4, 6) \rangle$ (Å)	$\langle d(4, 7) \rangle$ (Å)	$\langle d(4, 8) \rangle$ (Å)
α -helix	14.0 ± 0.9	9.3 ± 2.8	9.1 ± 1.7
3_{10} -helix	11.7 ± 1.8	8.8 ± 1.6	12.0 ± 1.2

The distances between side-chain β -carbons (Å) for the α -helix geometry ($\phi = -65^\circ$, $\psi = -40^\circ$) and the 3_{10} -helix geometry ($\phi = -74^\circ$, $\psi = -4^\circ$) as calculated using the program Insight II (BioSym) running on a personal IRIS computer (Silicon Graphics). For the ESR spectra of the doubly labelled peptides, $\delta^{-1}(\text{G}^{-1})$ is the reciprocal of the peak-to-peak width of the centre ($M_i=0$) nitroxide hyperfine line. This parameter is proportional to the distance between the side chains. Constrained molecular dynamics results for helices in the two helical geometries with labels at the indicated positions. The calculated average distance ($\pm$ s.d.) between nitroxide N atoms on each peptide is $\langle d(i, j) \rangle$. The 3K helix analogues with spin labels were built using InsightII version 1.1.0 (BioSym). Molecular dynamics calculations were performed using Discover v. 2.7 (BioSym) with a dielectric constant of 80. Side-chain atoms were unconstrained but helix backbone atoms were constrained using a strong restoring potential to maintain the helical structure. All of the structures were subjected to an initial 10 ps of molecular dynamics at 300 K followed by energy minimization using the steepest descent method. The structures were then heated to 900 K and molecular dynamics run for 20 ps. The temperature was then lowered to 300 K and molecular dynamics run for 20 ps to equilibrate the structures. After this annealing²¹, the calculations were run for 100 ps and analysed. This entire procedure was repeated, and yielded the same results to within experimental error.

line nitroxide spectrum ($\sim 25\%$ of the integrated area) and a very broad spectrum ($\sim 75\%$). The broad component shows that the average $d(4, 7)$ decreases with increasing helicity and further supports the assignment of a 3_{10} -helix structure. This superposition of two spectra may indicate the coexistence of the two helical forms, with α -helices giving rise to the narrow-line component.

We find that there is significant evidence that α -helix is not the correct structure for these Ala-based peptides. The data in this report can be readily explained by the presence of 3_{10} -helix geometry. Recent work with aminoisobutyric acid rich peptides provides a reference point for these findings. The amino acid is disubstituted at the α -carbon and this restricts its conformation to the helical region of the conformational energy map^{22,23}. In general, shorter peptides (eight residues or less) with a high aminoisobutyric acid content favour the formation of 3_{10} -helices. NMR investigations of 8-mers rich in aminoisobutyric acid in organic solvents have demonstrated that the $\alpha \rightleftharpoons 3_{10}$ equilibrium is readily controlled by the choice of solvent²⁴. Our finding is surprising as these Ala-based peptides are 16-mers that contain no aminoisobutyric acid. But we have studied aqueous solutions, in which the nature of the $\alpha \rightleftharpoons 3_{10}$ equilibrium is not well understood.

We consider two additional points. First, the $i+5$ placement of the lysine residues in the original 3K(I) was to ensure that these peptides would not be amphiphilic in the α -helix conformation. Coincidentally, this lack of amphiphilicity is preserved in the 3_{10} -helix conformation. Second, a recent molecular dynamics study of an α -helical fragment of ribonuclease A suggested the unfolding sequence followed: α -helix $\rightarrow 3_{10}$ -helix $\rightarrow$ no hydrogen-bond²⁵. We propose that short monomeric Ala-based peptides in aqueous solution actually coexist in both helical forms and that the population of the 3_{10} -helix conformation is much greater than previously expected. It may be that the 3_{10} conformation is mainly a long-lived intermediate between the two terminal states of α -helix and random coil. In contrast, peptides in aqueous solution may prefer the 3_{10} -helix conformation, whereas the α -conformation is favoured in folded proteins. $\square$

Received 29 June; accepted 25 August 1992.

- Marqusee, S., Robbins, V. H. & Baldwin, R. L. *Proc. natn. Acad. Sci. U.S.A.* **86**, 5286–5290 (1989).
- Merutka, G., Lipton, W., Shalongo, W., Park, S. H. & Stellwagen, E. *Biochemistry* **29**, 7511–7515 (1990).
- Chakrabarty, A., Schellman, J. A. & Baldwin, R. L. *Nature* **351**, 586–588 (1991).
- Padmanabhan, S., Marqusee, S., Ridgeway, T., Laue, T. M. & Baldwin, R. L. *Nature* **344**, 268–270 (1990).
- Manning, M. C. & Woody, R. W. *Biopolymers* **31**, 569–586 (1991).
- Osterhout, J. J. *et al. Biochemistry* **28**, 7059–7064 (1989).
- Bradley, E. K., Thomason, J. F., Cohen, F. E., Kosen, P. A. & Kuntz, I. D. *J. molec. Biol.* **215**, 607–622 (1990).
- Wüthrich, K. *NMR of Proteins and Nucleic Acids* (Wiley, New York, 1986).
- Voet, D. & Voet, J. G. *Biochemistry* (Wiley, New York, 1990).
- Barlow, D. J. & Thornton, J. M. *J. molec. Biol.* **201**, 601–619 (1988).
- Todd, A. P. & Millhauser, G. L. *Biochemistry* **30**, 5515–5523 (1991).
- Mick, S. M., Rodd, A. P. & Millhauser, G. L. *Biochemistry* **30**, 9498–9503 (1991).
- Byler, D. M. & Susi, H. *Biopolymers* **25**, 469–487 (1986).
- Kennedy, D. F., Crisma, M., Toniolo, C. & Chapman, D. *Biochemistry* **30**, 6541–6548 (1991).
- Malcolm, B. R. *Biopolymers* **22**, 319–321 (1983).
- Dwivedi, A. M., Krimm, S. & Malcolm, B. R. *Biopolymers* **23**, 2025–2065 (1984).
- Prestrelski, S. J., Byler, D. M. & Thompson, M. P. *Int. J. Peptide Protein Res.* **37**, 508–512 (1991).
- Luckhurst, G. R. in *Spin labeling: Theory and Applications* (ed. Berliner, L. J.) Ch. 4 (Academic, New York, 1976).
- Falle, H. R. *et al. Molec. Phys.* **11**, 49–56 (1966).
- Lemaire, H., Rassat, A., Rey, P. & Luckhurst, G. R. *Molec. Phys.* **14**, 441–447 (1968).
- Krystek, S. R. *et al. FEBS Lett.* **299**, 255–261 (1992).
- Toniolo, C. & Benedetti, E. *Trends. biochem. Sci.* **16**, 350–353 (1991).
- Karle, I. L. & Balaram, P. *Biochemistry* **29**, 6747–6756 (1990).
- Gautam, B., Bagchi, K. & Kuki, A. *Biopolymers* **31**, 1763–1774 (1991).
- Tirado-Rives, J. & Jorgensen, W. L. *Biochemistry* **30**, 3864–3871 (1991).

ACKNOWLEDGEMENTS. We thank R. L. Baldwin for discussion. This work was supported by grants from the N.S.F. and the N.I.H. and in part by the Petroleum Research Fund, administered by the American Chemical Society.

ERRATUM

Direct measurement of the optical depth in a spiral galaxy

Raymond E. White III & William C. Keel

Nature **359**, 129–131 (1992)

IN this letter in the 10 September issue, the third sentence in the bold first paragraph should be replaced with the following: "This statistically-derived result (concluding that spiral galaxies are opaque) is however vulnerable to several selection effects, and seems to contradict the fact that the (extragalactic) survey was successfully performed (since we live in a spiral galaxy which must then be partially transparent) as well as anecdotal examples of galaxies visible through other galaxies."

PLEASE follow these guidelines so that your manuscript may be handled expeditiously.

Nature is an international journal covering all the sciences. Contributors should therefore bear in mind those readers who work in other fields and those for whom English is a second language, and write clearly and simply, avoiding unnecessary technical terminology. Space in the journal is limited, making competition for publication severe. Brevity is highly valued. One printed page of *Nature*, without interruptions, contains about 1,300 words.

Manuscripts are selected for publication according to editorial assessment of their suitability and reports from independent referees. They can be sent to London or Washington and should be addressed to the Editor. Manuscripts may be dealt with in either office, depending on the subject matter, and will where necessary be sent between offices by overnight courier. High priority cannot be given to presubmission enquiries; in urgent cases they can be made in the form of a one-page fax. All manuscripts are acknowledged on receipt but fewer than half are sent for review. Those that are not reviewed are returned as rapidly as possible so that they may be submitted elsewhere without delay. Contributors may suggest reviewers; limited requests for the exclusion of specific reviewers are usually heeded. Manuscripts are usually sent to two or three reviewers, chosen for their expertise rather than their geographical location. Manuscripts accepted for publication are typeset from the London office.

Nature requests authors to deposit sequence and crystallographic data in the databases that exist for this purpose, and to mention availability of these data.

Once a manuscript is accepted for publication, contributors will receive proofs in about 4 weeks. *Nature's* staff will edit manuscripts with a view to brevity and clarity, so contributors should check proofs carefully. Manuscripts are generally published 2–3 weeks after receipt of corrected proofs. *Nature* does not exact page charges. Contributors receive a reprint order form with their proofs; reprint orders are processed after the manuscript is published and payment received.

Categories of paper

Review articles survey recent developments in a field. Most are commissioned but suggestions are welcome in the form of a one-page synopsis addressed to the Reviews Coordinator. Length is negotiable in advance.

Articles are research reports whose conclusions are of general interest and which are sufficiently rounded to be a substantial advance in understanding. They should not have more than 3,000 words of text (not including figure legends) or more than six display items and should not occupy more than five pages of *Nature*.

Articles start with a heading of 50–80 words written to advertise their content in general terms, to which editors will pay particular attention. The heading does not usually contain numbers, abbreviations or measurements. The introduction to the study is contained in the first two or three paragraphs of the article, which also briefly summarize its results and implications. Articles have fewer than 50 references and may contain a few short subheadings.

Letters are short reports of outstanding novel findings whose implications are general and important enough to be of interest to those outside the field. Letters should have 1,000 or fewer words of text and four or fewer display items. The first paragraph describes, in not more than 150 words and without the use of abbreviations, the background, rationale and chief conclusions of the study for the particular benefit of non-specialist readers. Letters do not have subheadings and should contain fewer than 30 references.

Commentary articles deal with issues in, or arising from, research that are also of interest to readers outside research. Some are commissioned but suggestions can be made to the Commentary Editor in the form of a one-page synopsis.

News and Views articles inform non-specialist readers about new scientific advances, sometimes in the form of a conference report. Most are commissioned but proposals can be made in advance to the News and Views Editor.

Scientific Correspondence is for discussion of topical scientific matters, including those published in *Nature*, and for miscellaneous contributions. Priority is given to letters of fewer than 500 words.

Preparation of manuscripts

All manuscripts should be typed, double-spaced, on one side of the paper only. An original and four copies are required, each accompanied by artwork. If photographs are included, five sets of originals are required; for line drawings, one set of originals and four good-quality photocopies are acceptable. Reference lists, figure legends and tables should all be on separate sheets, all of which should be double-spaced and numbered. Three copies of relevant manuscripts in press or submitted for publication elsewhere should be included with each copy of a submitted manuscript, and clearly marked as such. Five copies of revised and resubmitted manuscripts, labelled with their manuscript numbers are required, together with five copies of a letter detailing the changes made.

Titles are brief and simple. Active verbs, numerical values, abbreviations and punctuation are to be avoided. Titles should contain one or two key words for indexing purposes.

Artwork should be marked individually and clearly with the author's name and, when known, the manuscript number. Ideally, no figure should be larger than 28 by 22 cm. Figures with several parts are to be avoided and are permitted only if the parts are closely related, either experimentally or logically. Unlettered originals of photographs should be provided. Suggestions for cover illustrations, with captions and labelled with the manuscript number, are welcome. Original artwork is returned when a manuscript cannot be published.

Protein/nucleotide sequences should ideally be in the three-letter and not the single-letter code for amino acids. One column width of *Nature* can accommodate 20 amino acids or 60 base pairs.

Colour artwork. A charge of £500 per page is made as a contribution towards the cost of reproducing colour figures. Inability to pay these costs will not prevent publication of essential colour figures if the circumstances are explained. Proofs of colour artwork may be sent to contributors under separate cover from their galley proofs.

Figure legends should not exceed 300 words and ideally should be shorter. The figure is described first, then, briefly, the method. Reference to a method published elsewhere is preferable to a full description. Methods are not described in the text.

References are numbered sequentially as they appear in the text, followed by those in tables and finally by those in figure legends. Only papers published or in the press are numbered and included in the reference list. All other forms of reference should be cited in the text as a personal communication, manuscript submitted or in preparation. Text is not included in reference lists. References are abbreviated according to the *World List of Scientific Periodicals* (Butterworths, London, 1963–65). The first and last page numbers are included; reference to books should include publisher, place and date.

Abbreviations, symbols, units and greek letters should be identified the first time they are used. Acronyms should be avoided whenever possible and, if used, defined. Footnotes are not used in the text.

Acknowledgements are brief and appear after the reference list; grant and contribution numbers are not allowed.

Supplementary information is material relevant to Articles or Letters which cannot, for lack of space, be published in full, but which is available from *Nature* on request.

Submission. Manuscripts can be sent to the Editor at 4 Little Essex Street, London WC2R 3LF, UK or at 1234 National Press Building, Washington, DC 20045, USA. Manuscripts or proofs sent by air courier to London should be declared as 'manuscripts' and 'value \$5' to prevent the imposition of import duty and value-added tax.